

Niko Koski

JURIDINEN EUROOPANENGLANTI, COMMON LAW JA SUOMALAISTEN SÄÄDÖSTEN ENGLANNINKIELISET KÄÄNNÖKSET

Vertaileva korpustutkimus

Informaatioteknologian ja viestinnän tiedekunta
Pro gradu -tutkielma
kesäkuu 2019

TIIVISTELMÄ

Niko Koski: Juridinen euroopanenglanti, common law ja suomalaisten säädösten englanninkieliset käännökset : Vertaileva korpustutkimus
Pro gradu -tutkielma
Tampereen yliopisto
Monikielisen viestinnän ja käännöstieteen maisteriohjelma, englannin opintosuunta
kesäkuu 2019

Pro gradu -tutkielmassani pyrin selvittämään, miten mannereurooppalaisen kansallisen lainsäädännön epäviralliset englanninkieliset käännökset eroavat tekstuaalisilta piirteiltään Euroopan unionin lainsäädännön englanninkielisistä toisinoista sekä common law -oikeudellisesta kansallisesta lainsäädännöstä. Vertailevan korpusanalyysin lähtökohtana on niin kutsuttu käännösuniversaaliteoria, jota käytetään apuna tutkimusaineiston tarkastelussa.

Tutkimuksen aineistona on 26 suomalaisen lain epävirallista englanninkielistä käännöstä, 24 Euroopan unionin asetuksen englanninkielistä toisintoa sekä 18 Yhdistyneen kuningaskunnan parlamentin antamaa lakia. Tutkimusaineisto on koottu kolmeksi eri korpukseksi, joiden eroja tarkastellaan globaalisti makrotasolla (virkepituus, sanaluokat, TTR-luku, leksikaalinen tiheys) sekä paikallisesti mikrotasolla (sanasto, n-grammit, juridisen logiikan kommunikointi). Korpusanalyysissä saatuja havaintoja peilataan *eksplikoitumisen*, *normalisoitumisen* ja *yksinkertaistumisen* käännösuniversaalihypoteeseihin.

Tutkimuksen tarkoituksena on selvittää vertailevan korpustutkimuksen menetelmin, palautuuko suomalaisissa säädöskäännöksissä käytettävä englanti common law'hon vai Euroopan unionin instituutioissa muodostuneeseen juridiseen euroopanenglantiin. Hypoteesina on, että suomalaisten lakien englanninkieliset käännökset ovat tekstuaalisilta piirteiltään yhdenmukaisempia Euroopan unionin asetusten englanninkielisten toisintojen kanssa kuin Yhdistyneen kuningaskunnan parlamentin lakien kanssa.

Tutkielman teoreettisen viitekehyksen keskiössä on käännösuniversaaliteoria, jonka avulla tekstuaalisten piirteiden yhdenmukaisuutta korpusten välillä tarkastellaan. Oikeuskieltä ja oikeusjärjestelmiä tarkastellaan juridisten tekstien kääntämisen näkökulmasta, ja erityisesti juridisen englannin piirteitä tarkastellaan mannereurooppalaisten oikeusjärjestelmien ja common law -oikeuden näkökulmasta.

Tutkimuksessa saatiin selville, että suomalaisten lakien epävirallisten käännösten englanti on tekstuaalisilta piirteiltään lähempänä Euroopan unionissa käytettävää juridista englantia kuin common law -englantia. Tuloksista ei saatu tukea käännösuniversaalihypoteeseille, ja näyttää siltä, että institutionaaliset käytännöt sekä englannin kielen ongelmallinen asema eurooppalaisten juridisten tekstien kielenä vaikuttavat merkittävästi juridisissa teksteissä käytettävän englannin muotoon.

Avainsanat: juridinen euroopanenglanti, juridisten tekstien kääntäminen, käännösuniversaaliteoria, vertaileva korpustutkimus

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

Sisällysluettelo

1 Johdanto	1
2 Teoreettinen tausta	5
2.1 Oikeusjärjestelmät ja lainsäädännön laatimisen käytännöt	5
2.1.1 Common law ja mannereurooppalainen oikeus	6
2.1.2 Suomalainen lainsäädäntö	7
2.1.3 Yhdistyneen kuningaskunnan lainsäädäntö	8
2.1.4 Euroopan unionin lainsäädäntö	9
2.2 Juridinen englantia	10
2.2.1 Common law -englantia	10
2.2.2 Juridinen euroopanenglantia	12
2.3 Juridisten tekstien kääntämisen erityishaasteet	14
2.3.1 Oikeusjärjestelmien terminologinen vastaamattomuus	14
2.3.2 Oikeuskielen logiikan kääntäminen	17
2.4 Käännösuniversaalit	19
2.4.1 Käännösuniversaaliteorian kritiikkiä	19
2.4.2 Vertaileva korpusmenetelmä ja T-universaalit	20
3 Menetelmä ja aineisto	23
3.1 Tutkimuskysymys	25
3.2 Korpusmenetelmä ja korpustyökalut	25
3.3 Säädösten rakenteet	29
3.4 Suomalaiset lait	30
3.5 Yhdistyneen kuningaskunnan lait	31
3.6 Euroopan unionin asetukset	32
3.7 Aineiston syöttäminen korpuksiin	32
4 Korpusanalyysin tulokset	34
4.1 Tekstuaalinen yhdenmukaisuus makrotasolla	34
4.1.1 Virkepitäydät	34

4.1.2 Sanaluokat	37
4.1.3 Leksikaalinen tiheys	39
4.1.4 TTR-luku	41
4.2 Tekstuaalinen yhdenmukaisuus mikrotasolla	43
4.2.1 Sanalistat	43
4.2.2 N-grammifrekvenssit	47
4.2.3 Juridisen logiikan kommunikointi	49
4.3 Analyysin yhteenveto	53
4.3.1 Korpusten tekstuaalinen yhdenmukaisuus	53
4.3.2 Käännösuniversaalihypoteesit	54
5 Lopuksi	55
Lähdeluettelo	60
Liite 1: Lista aineistoon poimituista säädöksistä	64
Liite 2: Sanaluokkajakaumat korpuksissa	68
Liite 3: Verrokkikorpuksista ajetut sanalistat	69
Liite 4: Verrokkikorpuksista ajetut n-grammilistat	71
Summary in English	72

1 Johdanto

Kääntäminen on tärkeä osa juridista kielenkäyttöä alati globaalistuvassa maailmassamme, sillä informaatiota kuljetetaan oikeusjärjestelmien välillä käännösten kautta. Samalla oikeuskielisten tekstien kääntäminen kuitenkin poikkeaa monin tavoin muusta erikoisalakääntämisestä, sillä oikeuskieli on vahvasti sidoksissa yksittäisiin oikeusjärjestelmiin siinä missä useimmilla erikoisaloilla käsitejärjestelmät eivät niinkään poikkeaa siirryttäessä kulttuurista toiseen. Oikeutta käsitteellistetään lähtökohtaisesti eri tavoin eri oikeusjärjestelmissä, joten oikeusjärjestelmien välillä tekstejä käännettäessä törmätään terminologisen vastaamattomuuden ongelmaan. Eri oikeusjärjestelmissä oikeuskieli kietoutuu myös historiaan ja juridisen ammattikunnan käytäntöihin ja on muun muassa siksi yleensä vaikeaselkoista yleiskieleen verrattuna. Oikeuskielen tuominen lähemmäs tavallisen kielenkäyttäjän tarpeita onkin ollut ympäri englanninkielistä maailmaa 1970-luvulla perustetun Plain Legal English -liikkeen tarkoituksena (Mattila 2017, 425), ja maailmanlaajuisesti selkeämmän oikeuskielen asiaa ajaa kansainvälinen Clarity-järjestö (Williams 2014, 6).

Kansainvälisesti oikeuskommunikaatiota hankaloittaa myös se, että eri maissa oikeuskieli on muotoutunut vastaamaan eri oikeusjärjestelmien tarpeisiin. Erityisesti Englanti on hankala tapaus kansainvälisessä oikeuskommunikaatiossa, sillä useimmat englanninkielisten maiden oikeusjärjestelmät kuuluvat niin kutsuttuun *common law* -oikeuteen, joka eroaa hyvin paljon mannereurooppalaisesta oikeusperinteestä. Common law -Englanti tuottaa päänvaivaa mannereurooppalaiselle juristille erityisesti terminologisesta näkökulmasta. Kun Englantia käytetään *lingua francana* esimerkiksi sopimuksissa kahden mannereurooppalaisen oikeushenkilön välillä, joudutaan samalla törmäyskurssille common law -termistön kanssa, vaikka sopimusta tulkittaisiin tosiasiallisesti mannereurooppalaisten oikeusjärjestelmien näkökulmasta.

Tässä tutkielmassa pyrin selvittämään, missä määrin suomalainen juridisten tekstien käännösenglanti poikkeaa common law -oikeusjärjestelmien oikeusenglannista ja Euroopan unionissa muotoutuneesta juridisesta euroopanenglannista (engl. *EU legal English*). Suomalaisen juridisen käännösenglannin suhteutuminen common law -Englantiin ja juridiseen euroopanenglantiin on kiinnostavaa erityisesti oikeusjärjestelmien eroavuuden vuoksi; yhtäältä lainsäädännöllisten tekstien kuljettaminen common law'ta kuvaavan englannin kautta on ongelmallista käännösten kuranttiuden suhteen, ja toisaalta juridisen euroopanenglannin

luonteesta tiedetään hyvin vähän (juridista euroopanenglantia ovat tutkineet mm. Bajčić 2018 ja Robertson 2012).

Tutkimuksen menetelmänä on vertaileva korpustutkimus. Tutkimuksen aineistona on kolme osa-aineistoa, jotka koostuvat suomalaisten lakien englanninkielisistä käännöksistä, Euroopan unionin asetusten englanninkielisistä toisinnoista sekä Yhdistyneen kuningaskunnan parlamentin säätämistä laeista. Otos on koottu ympäristö- ja energiaoikeuden alan säädöksistä (otantaa kuvattu tarkemmin luvussa 3). Osa-aineistot kootaan kolmeen eri korpukseen, joiden eroja verrataan globaalisti makrotasolla (virkepituus, sanaluokat, TTR-luku (engl. *type-token ratio*), leksikaalinen tiheys) sekä paikallisesti makrotasolla (sanasto, n-grammit, juridisen logiikan kommunikointi).

Tutkimusaineiston korpusanalyysistä saatuja havaintoja peilataan niin kutsuttuihin *käännösuniversaalihypoteeseihin*, joiden mukaan käännöksillä on yleispiirteitä, jotka toistuvat kohdeteksteissä riippumatta lähde- ja kohdekielistä. Käännösuniversaalihypoteeseista tarkastellaan *eksplikoitumista*, *normalisoitumista* sekä *yksinkertaistumista*. Käännösuniversaalit auttavat hahmottamaan lainsäädännön käännösten tekstuaalisten piirteiden yhteensopivuutta kohdekielellä alun perin kirjoitettuihin lainsäädännöllisiin teksteihin. Käännösuniversaaliteorian ja eri käännösuniversaalihypoteesien pätevyydestä kiistellään (yleiskatsaus keskustelusta teoksessa Mauranen & Kujamäki 2004, 1–2), mutta käännösuniversaalihypoteesit toimivat hyvin eräänlaisena komparatiivisten korpusten mittatikkuna: niiden avulla voidaan tehdä johtopäätöksiä siitä, johtuvatko mahdolliset erot korpusten tekstuaalisissa piirteissä käännösprosessin sisäisistä vai ulkoisista piirteistä. Toisin sanoen, jos korpustutkimuksen tulokset ovat olennaisesti ristiriidassa käännösuniversaalihypoteesien kanssa, syitä eroavaisuuksiin löytynee todennäköisesti käännösprosessin ulkoisista tekijöistä, kuten tekstien institutionaalisten laatimiskäytäntöjen tai oikeusjärjestelmien tekstuaalisten konventioiden eroista.

Lähtökohtana tutkimuksessa on suomalaisten lakien englanninkielisten käännösten korpus, jonka tekstuaalisten piirteiden samankaltaisuuksia ja eroavaisuuksia verrataan Euroopan unionin asetusten englanninkielisiä toisintoja sisältävään korpukseen sekä Yhdistyneen kuningaskunnan parlamentin alun perin englanniksi antamia lakeja sisältävään korpukseen. Korpusanalyysillä pyritään vastaamaan kahteen tutkimuskysymykseen:

- 1) Miten käännetty kansallinen lainsäädäntö eroaa kohdekielisestä ylikansallisesta lainsäädännöstä ja kohdekielisestä kansallisesta lainsäädännöstä:
 - a. makrotasolla (virkepituus, sanaluokat, leksikaalinen tiheys, TTR-luku)?
 - b. mikrotasolla (sanasto, n-grammit, juridisen logiikan kommunikointi)?
- 2) Pätevätkö *eksplikoitumisen*, *normalisoitumisen* ja *yksinkertaistumisen* käännösuniversaalihypoteesit lainsäädännön käännöksissä?

Hypoteesina tutkimuksessa on, että suomalaisten lakien englanninkieliset käännökset ovat tekstuaalisilta piirteiltään yhdenmukaisempia Euroopan unionin asetusten englanninkielisten toisintojen kanssa kuin Yhdistyneen kuningaskunnan parlamentin antamien lakien kanssa, sillä vaikka myös suurin osa Euroopan unionin lainsäädännöllisistä teksteistä laaditaan alun perin englanniksi, pyritään Euroopan unionin tekstit laatimaan siten, että niiden kääntäminen kaikille unionin 24 viralliselle kielelle onnistuu. Täten oletan, että EU-tekstit ovat myös lähempänä käännöstekstejä kuin alun perin kohdekielellä kirjoitettuja lainsäädännöllisiä tekstejä.

Säädöskäännösten vertailevaa korpustutkimusta on tehty jonkin verran (esim. Biel 2014a vertailee puolalaisia lakeja ja puolankielisiä EU-asetusten toisintoja ja Piehl 2006 suomalaisia lakeja ja suomenkielisiä EU-asetusten toisintoja), ja juridisten tekstien kääntämistä on muutoinkin tutkittu runsaasti (esim. Šarčević 1997, Kinnunen 2006, Paunio 2011). Myös vertailevan oikeuslingvistiikan alalla, joka osaltaan liittyy olennaisesti myös juridisiin käännösteksteihin, on tutkimuksia suurin mitoin. Aiempaa tutkimusta suomalaisten säädösten epävirallisten englanninkielisten käännösten suhteutumisesta alun perin englanniksi kirjoitettuun lainsäädäntöön on kuitenkin julkaistu niukasti – jos ollenkaan. Englannin kielen merkitys myös suomalaisessa oikeuskäytännössä on kuitenkin kiistaton, sillä myös epävirallisia säädöskäännöksiä hyödynnetään ulkomaisissa tuomioistuimissa (Šarčević 1997, 277–278), ja englantia toimii usein yhteisenä kielenä sopimuksissa, kun osapuolet eivät jaa yhteistä kieltä. Juridisissa käännöksissä käytettävällä englannilla on siis kauaskantoisia vaikutuksia, ja tutkielmani tarjoaa täten tuoreen ja merkityksellisen näkökulman suomalaiseen juridiseen englantiin.

Korpusmenetelmien käyttö käännöstieteellisessä tutkimuksessa on kasvattanut suosiotaan 1990- ja 2000-luvuilla (Biel 2014a, 95). Käännöstieteellisen korpustutkimuksen voi jakaa kahtia paralleelikorpuksen tutkimukseen ja vertailevaan korpustutkimukseen, joista etummaisessa on kyse lähde–kohde-vertailusta ja jälkimmäisessä, jonka alaan myös tämä tutkimus lukeutuu, kohdetekstien vertailusta alun perin kohdekielellä kirjoitettuihin teksteihin

(mts. 92). Käännöstiede ei ole eksaktia, joten käännösten suhteita alun perin kohdekielellä kirjoitettuihin teksteihin on perinteisesti tutkittu pääsääntöisesti kvalitatiivisin tekstianalyysimenetelmin, mutta korpusten käyttö on avannut käännöstieteessä uuden tavan kvantifioida käännösten piirteitä. Vaikka korpusten avulla tehdyt havainnot käännöksistä ovat parhaimmillaankin vain yleistyksiä, mahdollistaa korpusmenetelmien käyttö kuitenkin tekstuaalisten piirteiden tarkastelun suuremmasta joukosta tekstejä kuin manuaalinen tekstianalyysi. Siten käännösten tutkiminen korpustyökalujen avulla saattaa tarjota suuntamerkkejä kvalitatiivisiin jatkotutkimuksiin, ja vertaileva korpustutkimus onkin vakiinnuttanut asemansa käännöstieteellisessä tutkimuksessa paralleelikorpusten tutkimuksen rinnalla (mts. 96).

Tutkielma rakentuu siten, että ensin esittelen tutkielman teoreettisen viitekehyksen. Taustoitan common law -oikeusjärjestelmän ja mannereurooppalaisen oikeusjärjestelmän eroja ja esittelen juridisissa teksteissä käytettävää englantia common law -englannin ja juridisen euroopanenglannin näkökulmasta. Avaan lisäksi juridisten tekstien kääntämisen erityishaasteita terminologian ja oikeuskielen logiikan näkökulmasta ja esittelen korpusanalyysin apuna käytettävät käännösuniversaalihypoteesit. Kolmannessa luvussa spesifioin tutkimuskysymykset sekä esittelen tutkimuksessa käytettävän korpusmenetelmän ja tarkasteltavan korpusaineiston. Neljännessä luvussa raportoin aineiston analyysin tulokset. Tutkielmani viidennessä, viimeisessä luvussa kokoon havaintoni yhteenvedoksi ja pohdin, miten tutkimukseni tulokset istuvat linjaan aiemmin juridisten tekstien kääntämisestä tehdyn tutkimuksen kanssa.

2 Teoreettinen tausta

Tässä luvussa esittelen tutkielman teoreettisen taustan. Ensin taustoitan common law -oikeuden ja mannereurooppalaisen oikeusjärjestelmän eroja. Tämä on tärkeä näkökulma tutkielmassani sen vuoksi, että tutkimuksen eri osa-aineistojen taustalla toimivat oikeusjärjestelmät vaikuttavat myös oikeuskieleen ja täten hypoteesini rakentumiseen. Käyn läpi myös esimerkiksi lainlaatumiskäytäntöjä ja oikeudenkäytön periaatteita Suomessa, Yhdistyneessä kuningaskunnassa sekä Euroopan unionissa. Sen jälkeen erittelen tarkemmin juridisen englannin historiaa ja erityispiirteitä common law -englannin ja juridisen euroopanenglannin valossa, sillä englannin kielen käyttöön mannereurooppalaisten oikeusjärjestysten kuvaamisessa liittyy monia ongelmia johtuen oikeusjärjestelmien keskinäisistä eroista. Tutkielman viitekehyksessä on syytä lisäksi avata juridisten tekstien kääntämisen erityishaasteita terminologian ja oikeuskielen logiikan näkökulmasta. Luvun lopuksi taustoitan niin kutsuttua käännösuniversaalioteoriaa ja esittelen käännösuniversaalihypoteesit, jotka toimivat apuvälineenä korpustutkimuksen tuloksia analysoitaessa.

2.1 Oikeusjärjestelmät ja lainsäädännön laatimisen käytännöt

Oikeusjärjestelmät toimivat monin erilaisin tavoin eri puolilla maailmaa, mutta oikeuskulttuurit voidaan jakaa karkeasti kahteen eri globaalisti merkittävään perheeseen: roomaanis-germaaniseen (mannereurooppalaiseen, engl. *civil law*) oikeuteen ja *common law* -oikeuteen (Mattila 2017, 166). Myös esimerkiksi islamilainen ja hindulainen oikeus vaikuttavat eräissä Afrikan ja Aasian maissa, mutta Heikki Mattilan mukaan oikeusjärjestysten käsitejärjestelmät vastaavat suurimmaksi osaksi myös näillä alueilla joko common law'ta tai mannereurooppalaista oikeutta, mitä selittää siirtomaa-aika ja länsimaalaistuminen (Mattila 2017, 166–167). Myös säädösten laatimisen käytännöt vaihtelevat huomattavasti eri oikeusjärjestelmien välillä. Suomalaisen oikeusjärjestelmän pohjautuessa mannereurooppalaiseen oikeuteen, englantilaisen common law'hon ja Euroopan unionin oikeuden sisältäessä piirteitä molemmista oikeusperinteistä on hyvä ensinnäkin määritellä tarkemmin common law -oikeuden ja mannereurooppalaisen oikeuden erot ja toisekseen kuvata lainsäädännön laatimisen eroja tutkielmassa käsiteltävien säädösten taustalla olevien oikeusjärjestelmien välillä.

2.1.1 Common law ja mannereurooppalainen oikeus

Common law tarkoittaa englantilaisperäistä oikeusperinnettä, joka on muodostunut ennakkotapausten eli *prejudikaattien* kautta. Common law:ssa tuomioistuimien päätökset ovat samalla tavoin lakia kuin kirjoitettu eli *kodifioitu* laki. Alemmat tuomioistuimet ovat tällöin sidottuja noudattamaan ylempien tuomioistuimien aiemmin antamia tuomioita. (Husa 2013, 303–304.) Common law -oikeuden maita ovat esimerkiksi Yhdysvallat (pl. Louisiana), Kanada (pl. Quebec), Australia ja Yhdistynyt kuningaskunta (pl. Skotlanti) (Ángeles & Llopis 2017, 20–21).

Mannereurooppalaisessa oikeudessa taas etusijalla on kodifioitu laki, jolla eri oikeusaloja säännellään hyvin tarkasti, ja tuomioistuimet eivät voi antaa tuomioita, jotka eivät pohjautu kirjoitettuun lainsäädäntöön (Husa 2013, 303–304). Mannereurooppalaisen oikeuden maihin kuuluvat muun muassa Espanja, Saksa, Ranska, Italia, Ruotsi ja Suomi. Esimerkiksi ranskalaisesta oikeusperinteestä poiketen Suomessa ennakkotapauksilla on suuri painoarvo; vaikka korkeimman oikeuden aiemmat tuomiot eivät varsinaisesti sido esimerkiksi käräjäoikeuksia, tosiasiaassa ennakkotapaukset tai yksittäisen tapauksen ennakkotapauspotentiaali vaikuttavat huomattavasti esimerkiksi valituslupien myöntämiseen (Jokela 2005, 233). Myös Suomen perustuslakiin kirjattu ihmisten yhdenvertaisen kohtelun periaate edellyttää, että samasta oikeuskysymyksestä ei anneta eri tuomioistuimissa toisistaan poikkeavia ratkaisuja (mts. 206). Eri puolilla Eurooppaa käytännöt vaihtelevat siis paljon, ja kaiken kaikkiaan oikeuskulttuurit ovatkin Mattilan mukaan hitaasti yhdentymässä (2017, 167 & 458).

Mannereurooppalaisten oikeusjärjestyksien ja common law'n välillä on kuitenkin yhä huomattavia eroja muun muassa oikeuden jaottelussa, käsitteistössä ja oikeuslähteissä, ja nämä erot selittyvät Mattilan mukaan oikeusperheiden historiallisella taustalla: common law -oikeus on muotoutunut tapaoikeutena normannivalloituksen jälkeen Englannin kuninkaallisissa tuomioistuimissa, ja mannereurooppalainen oikeusperinne on muodostettu keskiajan eurooppalaisissa yliopistoissa roomalaisen oikeuden pohjalta (Mattila 2017, 167). Mannereurooppalainen oikeusperinne on siis teoreettisempi siinä missä common law kietoutuu käytännön tuomioistuintoimintaan.

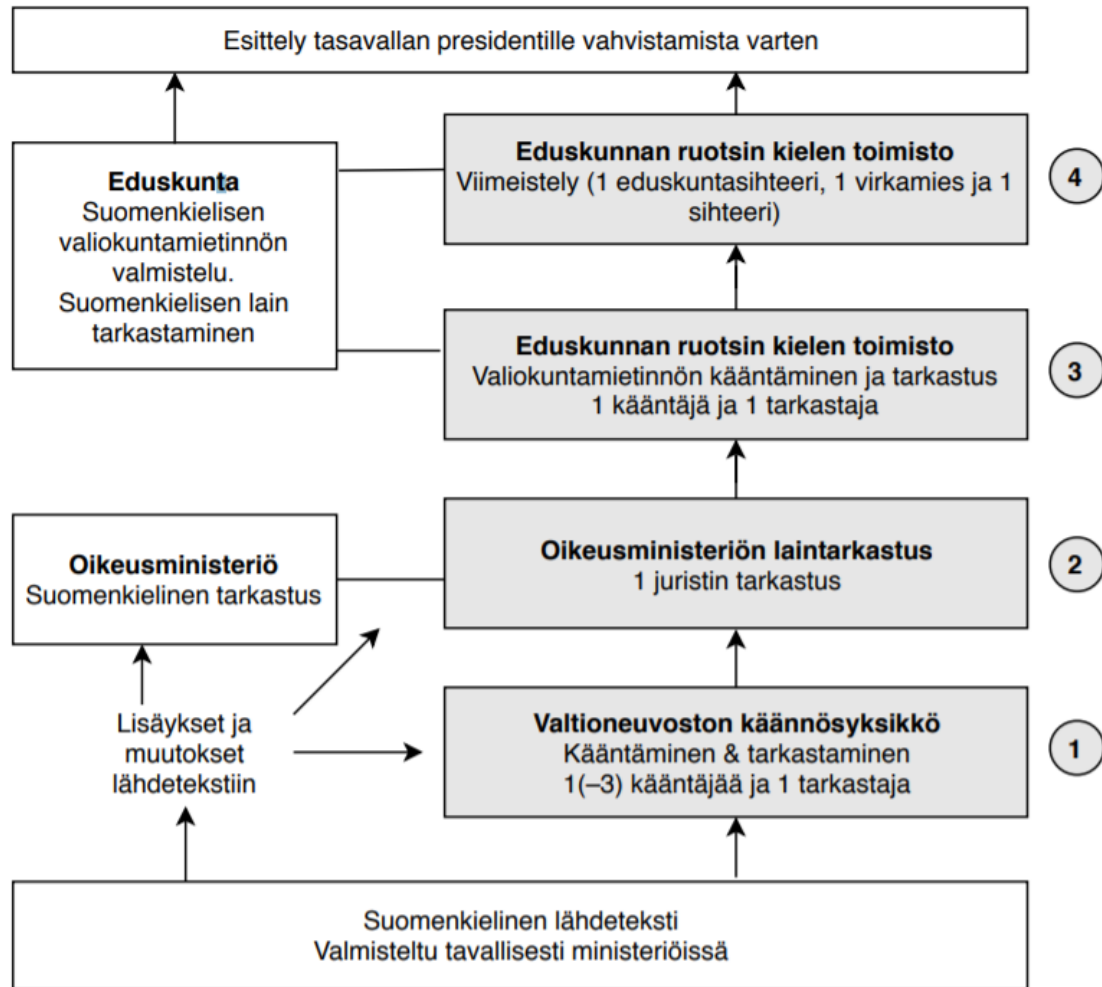
2.1.2 Suomalainen lainsäädäntö

Suomi kuuluu pohjoismaiseen oikeusperheeseen, joka on mannereurooppalaisen oikeusryhmittymän osa (Mattila 2017, 166). Oikeuslähdehierarkiassa Euroopan unionin asetukset ovat Suomen kansallisen lainsäädännön yläpuolella, eli mikäli Euroopan unionin lainsäädäntö on ristiriidassa suomalaisen säädöksen kanssa, sivuutetaan ristiriidassa oleva kansallinen säännös, toisin sanoen oikeusnormin sisältävä lainkohta, tuomioistuimessa ja noudatetaan Euroopan unionin asetusta (Jäsenvaltion oikeus – Suomi).

Suomella on kaksi kansalliskieltä, suomi ja ruotsi, joilla kaikki lainsäädäntö on todistusvoimaista (lainsäädännössä käytettävistä kielistä säädetään kielilain 30 §:ssä). Ruotsinkielinen lainsäädäntö Suomessa on käännetty suomenkielisten versioiden pohjalta (Lainvalmistelun prosessiopas), joten käännöstyö eroaa jonkin verran esimerkiksi Kanadan ja Euroopan unionin *co-drafting*-käytännöistä (kattavasti Kanadan järjestelmästä teoksessa Šarčević 1997).

Englanninkieliset versiot suomalaisesta lainsäädännöstä eivät ole virallisia eivätkä siten juridisesti sitovia, ja epävirallisiin käännöksiin onkin usein lisätty varaus ”NB: Unofficial translation. Legally binding only in Finnish and Swedish”. Käännöksen virallisuuden lisäksi olennainen ero ruotsin- ja englanninkielisten käännösten välillä on, että ruotsinkielisiä versioita luonnostellaan jo lainvalmistelun aikana (ks. kuva 1 alla), jolloin myös kääntäjät ovat mukana lainsäädäntöprosessissa.

Englanninkieliset käännökset taas tehdään eri ministeriöissä julkaistun lain pohjalta, eikä kääntäjä ole tällöin todennäköisesti ollut lainsäädäntöprosessissa mukana. Kääntäjällä saattaa olla mahdollisuus päästä käsiksi lainvalmistelupöytäkirjoihin, mutta käännöksiä ei ole tuotettu lainvalmistelun osana, jolloin alustavista käännöksistä ei ole esimerkiksi annettu lausuntoja eikä täten voida samalla tavalla varmistua käännöksen oikeellisuudesta. Susan Šarčevićin mukaan käännöksen tuotannon aikaiset tilanteelliset tekijät vaikuttavat suoraan sen tulkinnalliseen arvoon (Šarčević 1997, 95), ja kääntäjän liittäminen dynaamiseksi osaksi lainvalmisteluprosessia parantaa juridisten käännösten laatua (mts. 271–272). Suomalaisten lakien ja muiden säädösten englanninkieliset ja ruotsinkieliset käännökset ovat siis epäilemättä keskenään epäsymmetrisiä tässä suhteessa.



Kuva 1. Suomalaisen lain ruotsinkielisen version käännösprosessi (Nordman 2009, 45 [oma käännös])

2.1.3 Yhdistyneen kuningaskunnan lainsäädäntö

Yhdistyneen kuningaskunnan parlamentin antamat lait (engl. *Acts*) ovat koko kuningaskunnan yhteistä lainsäädäntöä. Yhdistynyt kuningaskunta on jaettu kolmeen oikeudenkäyttöalueeseen: Englanti ja Wales muodostavat yhden oikeudenkäyttöalueen, Skotlanti omansa ja Pohjois-Irlanti omansa. Näiden oikeudenkäyttöalueiden välillä tuomioistuinkäytännöt vaihtelevat, sillä Skotlannin ja Pohjois-Irlannin paikalliset oikeusperinteet kulkevat rinnakkain Yhdistyneen kuningaskunnan parlamentin säätämien lakien ja kansallisten common law -prosessien kanssa. Parlamentin antamat lait kuuluvat kuitenkin kaikkialla niin kutsuttuun primaarilainsäädäntöön. (Jäsenvaltion oikeus – Englanti ja Wales.) Koska tutkielmani koskee ainoastaan lainsäädäntöä eikä tuomioistuinkäytäntöjä, puhun tässä tutkielmassa yleisesti Yhdistyneen kuningaskunnan lainsäädännöstä riippumatta siitä, onko yksittäisten lakien vaikutuksesta jätetty pois joitain hallintoalueita, ja ottamatta huomioon oikeudenkäyttöalueiden eriäviä tuomioistuinkäytäntöjä.

Yhdistynyt kuningaskunta on vielä tätä kirjoittaessa Euroopan unionin jäsen, joten Euroopan unionin lainsäädäntö on oikeuslähdehierarkiassa Yhdistyneen kuningaskunnan lainsäädännön yläpuolella, eli mikäli Euroopan unionin lainsäädäntö on ristiriidassa Yhdistyneen kuningaskunnan lain kanssa, sivuutetaan ristiriidassa oleva kansallinen säännös, eli oikeusnormin sisältävä lainkohta, tuomioistuimessa ja noudatetaan Euroopan unionin asetusta (Jäsenvaltion oikeus – Englanti ja Wales).

2.1.4 Euroopan unionin lainsäädäntö

Euroopan unionissa on kirjoitushetkellä 28 jäsenmaata. Unioni antaa erilaisia määräyksiä, joilla säännellään jäsenmaiden lainsäädäntöä: *asetuksia, direktiivejä, päätöksiä ja suosituksia*. Näistä *asetukset* ovat varsinaista lainsäädäntöä ja ne tulevat sellaisenaan voimaan jäsenmaissa. Asetukset ovat oikeuslähdehierarkiassa jäsenmaiden kansallisen lainsäädännön yläpuolella, joten kansalliset säädökset, toisin sanoen lait ja asetukset, yleensä muokataan yhteensopiviksi Euroopan unionin asetusten kanssa. *Direktiivit* ovat ohjeistuksia, joiden pohjalta jäsenmaiden on laadittava omat säädöksensä. *Päätöksillä* täydennetään asetuksia ja direktiivejä, ja vapaaehtoiset *suositukset* eivät velvoita jäsenmaita. (Euroopan unionin toimintaperiaatteet.)

Euroopan unionin tuomioistuimessa kaikki unionin viralliset kielet, joita on tällä hetkellä 24, ovat todistusvoimaisia (Euroopan unionista tehdyn sopimuksen konsolidoitu toisinto, 55 artiklan 1 kohta). Tämä korostaa sen tärkeyttä, että EU-oikeudessa käytettävä monikielinen terminologia on johdonmukaista ja oikein leksikografisin perustein muodostettua. Mattilan mukaan Euroopan unionin määräykset laaditaan samanaikaisesti usealla kielellä, yleensä komission kolmella työkielellä, joita ovat englanti, ranska ja saksa (Mattila 2017, 41). Šarčević käyttää tästä monikielisestä lainsäädännön laatimisesta nimitystä *co-drafting* (Šarčević 1997, 99). Sen sijaan Agnieszka Doczekalskan (2018, 105) mukaan useimmiten alkuperäinen luonnostelukieli EU:n määräyksissä on englanti, josta tekstit käännetään muille jäsenvaltioiden kielille. Yksittäisten Euroopan unionin direktiivien ja asetusten englanninkielisten toisintojen ei voi siis varmuudella sanoa olevan alun perin englanniksi kirjoitettuja tai käännöksiä, sillä tätä tietoa ei ole dokumentoitu julkisesti, mutta esimerkiksi vuonna 2008 72,5 prosenttia kaikista Euroopan komission teksteistä laadittiin alun perin englanniksi ja vain 11,8 prosenttia ranskaksi (Felici 2015, 126). On myös arvioitu, että EU:n monikielisestä lainsäädännöstä jopa 95 prosenttia laaditaan alun perin englanniksi (Baaij 2012, 19). Koska englanti on joka tapauksessa oleellinen kieli Euroopan unionin lainsäädännön monikielisessä laatimisessa ja

selvästi suurin osa EU-lainsäädännöstä laaditaan alun perin englanniksi, voidaan Euroopan unionin englanninkieliseen lainsäädäntöön suhtautua alun perin lähdekielellä kirjoitettuna.

Euroopan unionin oikeusjärjestelmä on perustaltaan alun perin mannereurooppalainen, sillä unionin alkuperäiset perustajajäsenet kuuluvat mannereurooppalaiseen oikeusryhmittymään, mutta Yhdistyneen kuningaskunnan liittyttyä Euroopan unioniin vuonna 1973 myös common law -perinteet ovat vaikuttaneet unionin oikeusjärjestelmään (Mattila 2017, 167–168). Mattilan mukaan EU:n tuomioistuimen oikeudenkäyntimenettelyihin on Ison-Britannian vaikutuksesta muotoutunut common law -prosessien piirteitä sekä eräänlainen ennakkotapausjärjestelmä, joten Euroopan unionin oikeusjärjestelmän ei voi enää katsoa olevan täydellisesti joko mannereurooppalaisen oikeuden tai common law'n mukainen, vaan se on eräänlainen hybridijärjestelmä (mts. 168).

2.2 Juridinen englantia

Oikeuskielinen englantia poikkeaa hyvin paljon yleiskielisestä englannista niin sanastoltaan, kieliopiltaan kuin jopa ääntämyseltäänkin. Ranskan ja latinan vaikutus englannin kieleen korostuu oikeuskielessä aivan erityisellä tavalla. Juridinen englantia onkin muotoutunut kuvaamaan englantilaisperäistä common law -oikeusperinnettä, minkä vuoksi englannin käyttö mannereurooppalaisten oikeusjärjestysten kuvaamiseen on ollut perinteisesti ongelmallista. Yhdistyneen kuningaskunnan liittyttyä Euroopan unioniin on englantia kuitenkin syrjäyttänyt ranskan kielen valta-aseman unionin sisällä, ja on väitetty, että Euroopan unionin mannereurooppalaiseen oikeusperinteeseen pohjautuvaa oikeusjärjestelmää kuvaamaan on kehittynyt uudenlainen juridisen englannin variantti. Ero common law -oikeusjärjestelmien englantia on päällisin puolin erityisesti terminologinen, kun EU-instituutioille ja -käsitteille on jouduttu keksimään common law -järjestelmistä erilliset termit, mutta common law -englannin ja juridisen euroopanenglannin erot saattavat liittyä myös laajemmin kielen syntaktisiin ja tekstuaalisiin piirteisiin, sillä juridisia tekstejä tulkitaan eri tavoin common law -oikeudessa ja mannereurooppalaisessa oikeudessa.

2.2.1 Common law -englantia

Common law -oikeusjärjestelmien englannin tekee erityisen mielenkiintoiseksi ja hankalaksi sen koostuminen niin monenlaisista osista. Common law -englannin kehittymiseen ovat vaikuttaneet erityisesti latina ja juridinen angloranska eli *Law French*. Normannivalloituksen jälkeisenä aikana Englannissa yläluokka puhui ranskaa, mutta oikeuskielenä oli latina. Vaikka

1200-luvulla ranska alkoi hävitä Englannissa käyttökielenä, syrjäytti se latinan oikeuskielenä, ja 1500-luvulle asti englantilaisissa tuomioistuimissa käytettiin *Law French* -oikeuskieltä. (Mattila 2017, 407.) Tosin latinaksi kirjoitettiin vielä 1700-luvullekin asti esimerkiksi haasteita ja tuomioistuinten pöytäkirjoja, sillä latinaksi oli olemassa vakiintunut termistö (mts. 408). Näiden molempien kielten rinnalla kulki tietysti kansankieli englantia, jota kuultiin jo esimerkiksi parlamentissa 1300-luvulla (mp.). Mattila kuvaa common law'n juridista englantia termillä ”juridinen triglossia”, sillä kaikilla kolmella kielellä on ollut olennainen vaikutus nykyaikaisen oikeusenglannin muotoutumiseen (mts. 409). *Law French* -kielen vaikutuksen voi tätä nykyä nähdä muun muassa sanastossa (esim. *bailiff*, *attorney*, *misdemeanor*, *verdict*), sananmuodostuksessa (esim. *employer/employee*, *mortgagor/mortgagee*) sekä kiteytyneiden ilmauksien sanajärjestyksessä (adjektiivi substantiivin jäljessä joissain kiteytyneissä ilmauksissa, esim. *accounts payable*, *fee simple*) (Mattila 2017, 415). Latinasta sen sijaan tulee englantilaiseen oikeuskieleen paljon lyhenteitä, kuten *ss.* ja *e.g.*, sekä kiteytyneitä maksiimeja kuten *de minimis non curat lex* ja *pro bono*.

Common law -kieleen vaikuttaa paitsi edellä mainittu juridinen triglossia myös oikeusperinteen historia itsessään. Ennakkotapausten kautta muodostuneessa kuninkaan oikeudessa säädökset olivat vain suuntaviivoja; säädöstä ei edelleenkään pidetä täydellisenä ennen kuin tuomarit tarkentavat ennakkopäätöksissään, miten sitä tulee tulkita (Ángeles & Llopis 2017, 25). Common law -oikeusenglanti onkin tästä johtuen kirjoitettuna yleisluontoista, jotta teksti kattaisi mahdollisimman monenlaiset tilanteet, mutta toisaalta yksiselitteisen tarkkaa, jotta osapuolten oikeusasema olisi riidaton (Mattila 2017, 422). Siinä missä mannereurooppalaisten oikeusjärjestelmien juristit tulkitsevat María Ángelesin ja Orts Llopisin (2017, 24) mukaan lainsäädäntöä teleologisesti eli ottaen tulkinnassa huomioon lainsäätäjän alkuperäisen tarkoituksen tekstin laatimiselle, punnitsevat common law -juristit tarkoin jokaisen sanan mahdollista merkitystä, mikä taas johtuu siitä, että common law -perinteessä teksteistä pyritään tekemään mahdollisimman tarkkoja mutta toisaalta joustavia, jotta voidaan varmistua siitä, että teksti kattaa joka ikisen laatimishetkellä ennustamattomankin tilanteen (mp.). Tästä johtuen common law'n kieli on myös erittäin monisanaista (Mattila 2017, 422). Oikeusriitojen aloittamisen kynnys on matala ja ennakkotapaukset toimivat oikeuslähteenä, minkä vuoksi esimerkiksi sopimuksissa porsaanreikiä, joita vastapuoli voisi mahdollisesti hyödyntää, tukitaan erilaisilla varauksilla kuten *including, but without limiting the generality of the foregoing* (mp.). Monisanaisuuden liittyy myös toistavat sanaparit, joissa on käytännössä synonyyminen anglosaksinen ja latinalais-ranskalainen sana, kuten *acknowledge and confess*,

devise and bequeath tai *goods and chattels*. Mattilan mukaan tällä toistolla pyritään varmistamaan tekstin ymmärrettävyyttä ja yksiselitteisyyttä. (Mts. 417.)

2.2.2 Juridinen euroopanenglanti

Englantia käytetään juridisena *lingua francana* käytännössä kaikkialla Manner-Euroopassa sekä mannereurooppalaisten maiden ja common law -maiden välisessä kaupankäynnissä (Ángeles & Llopis 2017, 19). Lisäksi sitä käytetään usein linkkikielenä harvinaisempien kielten välillä erityisesti uudemmissa oikeudenaloilla, joilla on vakiintunut englanninkielinen termistö (Mattila 2017, 44–45). Mattilan mukaan kansainvälisissä sopimuksissa kieltä pyritään usein käyttämään samoin kuin common law -maissa, sillä mailla ei useinkaan ole olemassa riidatonta normipohjaa tai yhteistä kieltä, jolloin on järkevää tehdä sopimus common law’n tapaan mahdollisimman yksityiskohtaiseksi ja varmaksi (mts. 436). Juridinen englanti kuitenkin poikkeaa huomattavasti muusta englannista, ja common law -oikeusjärjestelmiin pohjautuvasta juridisesta englannista koituu usein väärinkäsityksiä ja hämmennystä, koska mannereurooppalaiset kääntäjät ja juristit eivät tunne sen erityispiirteitä. Mattila antaa esimerkiksi sanan *equitable*, joka Manner-Euroopassa tarkoittaisi *kohtuullista*, mutta common law -maissa ”*equity*-järjestelmän mukaista” (mts. 437). Euroopan unionissa on ongelmien ratkaisemiseksi kehitetty termistöä, joka ei olisi sidoksissa minkään yksittäisen maan oikeusjärjestelmään (Mattila 2017, 441; Šarčević 2018, 12). Laaja englannin käyttö mannereurooppalaisissa oikeusinstituutioissa vaatii, että kieleen luodaan uudissanoja, jotka kuvaavat esimerkiksi EU-instituutioita ja eri maiden oikeusinstituutioita ja -käsitteitä (Mattila 2017, 441).

Common law ei kuitenkaan eroa mannereurooppalaisesta oikeudesta pelkästään termistöltään. Ángelesin ja Llopisin kuvaama common law -juristien pakkomielle tekstin tarkkaan noudattamiseen (2017, 24) on yksi syy, miksi englannin käyttö juridisena *lingua francana* on ongelmallista. Heidän mukaansa mannereurooppalaisten oikeusjärjestelmien lainsäädännölliset tekstit rakentuvat yksinkertaisista virkkeistä, joissa todetaan laveasti säädöksen linjat ja periaatteet sen sijaan, että joka ikinen mahdollisuus otettaisiin huomioon kuten common law -oikeusjärjestelmän juridisissa teksteissä (mp.).

Tutkijat ovat kuitenkin peräänkuuluttaneet neutraalia juridisen englannin varianttia kansainvälisen oikeuskommunikaation tarpeisiin (esim. Ángeles & Llopis 2017, 27; Šarčević 2018, 12). Ángelesin ja Llopisin mukaan common law’sta erillistä juridisen englannin varianttia tarvitaan, jotta juridisten tekstien laatiminen olisi helpompaa ei-natiivipuhujille ja jotta common

law -englannin jäykkä tyyli on uskollisuus korvautuisi tehokkaalla viestinnällä (Ángeles & Llopis 2017, 27).

Bajčić (2018), Mattila (2017) ja Robertson (2012) puhuvatkin jo uudesta oikeusenglannin variantista, juridisesta euroopanenglannista (engl. *EU legal English*). Siihen nähden, että common law -oikeudessa juridisia tekstejä ylipäänsä tulkitaan erilaisista lähtökohdista kuin Manner-Euroopassa, Euroopan unionin instituutioissa tuotettavien englanninkielisten juridisten tekstien syntaktisia ja tekstuaalisia piirteitä on kuitenkin tutkittu kovin vähän. Colin Robertson (2012) on verrannut Euroopan unionin direktiivien kieltä Yhdistyneen kuningaskunnan sekä Ranskan lakien kieleen, ja todennut EU-englannin kieliasun olevan lähempänä ranskalaista perinnettä kuin common law -englantia (mts., 1233), mutta toteaa kyseessä olevan pikemmin uusi oikeuskielen genre kuin kuuluvan selkeästi common law'hon tai mannereurooppalaiseen oikeuteen (mts., 1237). Łucja Biel havaitsi tutkimuksessaan (2015, 145), että englanninkielisissä EU-teksteissä käytetään prepositiorakenteita huomattavasti enemmän kuin Yhdistyneen kuningaskunnan lainsäädännössä, mikä saattaa kieliä paitsi eksplikoitumisesta myös nominalisaatiosta ja pitkistä substantiivirakenteista (mp.).

EU-instituutioiden juridinen euroopanenglanti ei joka tapauksessa muistuta common law -perinteen oikeusenglantia, sillä EU-englanti on muotoutunut käännösten kautta, lähinnä ranskasta (Robertson 2012, 1237; Biel 2015, 150). Euroopan unionissa tekstit myös laaditaan käännettäviksi, millä on todennäköinen vaikutus käytettävään kieleen; Biel sanoo (mp.), että koska EU-englanti taipuu niiden kielten tarpeisiin, joihin siitä käännetään, on EU-englanti oma juridisen englannin varianttinsa. Suuri osa englanninkielisistä EU-teksteistä on myös ei-natiivipuhujien laatimia (Felici 2015, 124), joten laatijoiden äidinkielen interferenssi on mahdollinen EU-englannissa. Felicin mukaan juridiseen euroopanenglantiin toistaiseksi vielä vaikuttavat laadittavan tekstin oikeusala, laatijoiden kansallisuus ja englannin kielen taito sekä jäsenmaiden poliittiset intressit, eikä juridiselle euroopanenglannille ole vielä yhtenäistä laatimistyyliä (mts. 129). Felicin mukaan EU-tekstien syntaktiset piirteet ja rakenne vaikuttavat myös EU-kieleen, mutta niitä ei ole tutkittu samalla laajuudella kuin käsitteitä ja termejä (mts. 131).

2.3 Juridisten tekstien kääntämisen erityishaasteet

Juridisten tekstien kääntämisen voi katsoa kuuluvan erikoisalan kääntämiseksi (engl. *LSP translation*, *LSP* = *language for specific purposes*). Juridisten tekstien kääntäminen kuitenkin poikkeaa huomattavasti muiden erikoisalojen kääntämisestä, sillä siinä missä luonnontieteiden ja muiden eksaktien tieteiden alalla termit ovat usein toisiaan vastaavia kulttuurista riippumatta, käsitteellistetään oikeutta lähtökohtaisesti eri tavalla eri oikeusjärjestyksissä (Šarčević 1997, 66–67). Bielin mukaan juridisten tekstien kääntämisessä muuhun kääntämiseen verrattuna onkin lisähaasteena oikeustermien ja käsitejärjestelmien vastaamattomuus, joka johtuu oikeusjärjestelmien ja tulkintasääntöjen eroista eikä tarkoita ainoastaan yksittäisten termien käsite-eroja, vaan kokonaiset käsitejärjestelmät saattaa olla rakennettu toisistaan poikkeavasti (Biel 2014a, 50). Kääntäjän kannalta haasteellisena erityispiirteenä juridisissa teksteissä on lisäksi juridinen logiikka, jonka kommunikoinnin kääntäjän tulee hallita kyetäkseen tunnistamaan oikeussääntöjen merkityssuhteet ja tuottamaan kohdesäännöksen, jolla on sama juridinen vaikutus kuin lähdesäännöksellä (Šarčević 1997, 112).

2.3.1 Oikeusjärjestelmien terminologinen vastaamattomuus

Eri oikeuskulttuureissa käsitejärjestelmät ovat rakentuneet aikojen saatossa erilaisiksi, ja oikeusjärjestelmien käsite-erot aiheuttavat kääntäjille väärinkäsityksiä oikeusjärjestelmästä toiseen siirryttäessä. Vaikka eri oikeuskulttuureissa puhuttaisiin samalla termillä päällisin puolin samasta käsitteestä, esimerkiksi *avioliitosta*, on oikeuskulttuurien välillä kuitenkin erilaisia perinteitä ja eroja sen suhteen, mitä kaikkea termin käsitteisisältöön kuuluu (Bogdan 1993, 53). Lisäksi usein rajanveto yleiskielisen ilmauksen ja juridisen termin välillä, esimerkiksi juuri *avioliiton* tapauksessa, on hankalaa. Elina Paunio (2011, 29) huomauttaa, että kieli heijastuu ajatteluun ja että myös saman kielen jakavat ihmiset käsitteellistävät maailmaa toisistaan poikkeavasti. Esimerkiksi termi *domicile* (suom. lähivastine *kotipaikka*) käsitetään eri laajuiseksi amerikkalaisessa ja englantilaisessa oikeustermistössä (Šarčević 1997, 232). Mattila kiteyttää hyvin juridisen kääntämisen ongelmallisuuden tästä näkökulmasta:

[J]uridisten käsitejärjestelmien ja argumentaatiomenetelmien eroavuudet aiheuttavat helposti väärinkäsityksiä oikeustermistön suhteen – jopa silloin, kun kyse on kommunikaatiosta saman luonnollisen kielen puitteissa. Jos sama termi tarkoittaa yhdessä oikeuskulttuurissa yhtä ja toisessa toista, se ei välitä viestiä oikein ilman lisäselityksiä. (Mattila 2017, 454.)

Juridisen kääntämisen tutkimuksessa tulisikin kiinnittää ensisijainen huomio *tulkintaan*. Termejä tulkitaan eri oikeuskulttuureissa erilaisista lähtökohdista, ja oikeuskielisten tekstien kääntämisen optimistisena pyrkimyksenä on tuottaa kohdeoikeuskulttuurissa

lähdeoikeuskulttuuriin nähden ekvivalentti tulkinta. Šarčević (1997, 112) puhuu erityisesti kanadalaisiin säädöskäännösten laatimiskäytäntöihin viitaten ”kääntäjän vapautumisesta”, jossa on siirrytty mekaanisesta sana sanalta kääntämisestä siihen, että lainsäädännön kääntämisessä tärkeintä on tuottaa lähdetekstiin nähden ekvivalentti juridinen vaikutus (engl. *equivalence in legal effect*).

Lainsäädännöllisten tekstien kääntäjien on siis kyettävä arvioimaan, mikä on ollut tekstin tarkoituksena ja miten tuomioistuimet tulkitsevat lähdetekstiä sekä miten varmistetaan se, että kohdetekstin tulkinta ei poikkeaisi lähdetekstin tulkinnasta. Tulkinnan ekvivalenssin varmistamista vaikeuttavat muun muassa oikeuskulttuurien välinen terminologinen vastaamattomuus eli erot käsitteistöissä eri oikeusjärjestelmien välillä (Šarčević 1997, 121). Elina Paunio huomauttaa, että käytetystä termivastineesta riippumatta juridinen vaikutus ei välttämättä ole yhdenmukainen lähdesäännöksen kanssa (Paunio 2011, 29). Mattila kiteyttää: ”Kääntäjän on kyettävä välittämään lähdekielen taustalla olevan oikeusjärjestelmän elementit oikeasuhtaisina henkilöille, joille järjestelmä on kokonaan vieras” (Mattila 2008, 149). Kääntäjän onkin siis tärkeää pitää mielessä tulkinnan viitekehys ja se, minkälaisia tulkintaeroja kohdeoikeuskulttuurissa voi ilmetä lähdeoikeuskulttuuriin nähden.

Epävirallisten säädöskäännösten kohdalla vastuu tulkinnasta voi kuitenkin olla myös lukijalla. Tuija Kinnusen mukaan lakitekstien käännöksiä lukevien tuomareiden oletetaan tuntevan ennalta vieraan oikeusjärjestyksen peruskäsitteet, ja lakitekstin käännöstä pitäisi siksi tarkastella lähdeoikeusjärjestykseen kuuluvana vieraskielisenä tekstinä, ei kohdekulttuurisena tekstinä (Kinnunen 2006, 21). Kinnusen huomio ei oletettavasti päde virallisiin lainsäädännön käännöksiin, esimerkiksi suomalaisten säädösten ruotsinkielisiin käännöksiin. Toisaalta Anja Krogsgaard Vesteragerin mukaan todistusvoimaisissa käännöksissä suositaan usein edelleen sananmukaista kääntämistä, kun taas funktioltaan informatiiviset epäviralliset käännökset ovat vapaampia (Krogsgaard Vesterager 2017, 107). Šarčević taas ei tee eroa virallisten ja epävirallisten käännösten välille, vaan kääntäjän olisi kyettävä tunnistamaan lähdesäännöksen normatiivinen sisältö ja muotoilemaan kohdetekstissä säännös, jolla on sama oikeudellinen vaikutus (Šarčević 1997, 137). Kääntäjän rooli juridisten tekstien kääntämisessä nähdään siis monin eri tavoin riippuen käännetyn tekstin funktiosta. Mikäli kääntäjälle on kuitenkin annettu vapaat kädet muotoilla kohdesäännös lähdesäännöksen pohjalta, kääntäjälle on erilaisia strategioita oikeusjärjestelmien välisen terminologisen vastaamattomuuden ratkaisemiseksi.

Näihin strategioihin kuuluvat esimerkiksi lähivastine, luetteleva käännös, selittävä parafraasi, uudistermi ja sitaattilaina).

Mattilan mukaan oikeussäännöt poikkeavat toisistaan eri oikeusjärjestelmissä lähestulkoon aina jollain tavalla, kun mennään tarpeeksi syvälle yksityiskohtiin, joten kääntämisessä on kyse lähinnä *lähivastineiden* käytöstä (Mattila 2008, 152). Toisin sanoen *dynaamisen ekvivalentin*, eli juridisessa kontekstissa termivastineen, jonka käsitesisältö on sama kuin lähdetermillä, löytäminen ja käyttäminen vaatii, että lähde- ja kohdeoikeusjärjestelmien käsitteistöt ovat keskenään yhteensopivia (Šarčević 1997, 236). Lähivastineilla on Mattilan mukaan oltava yhteinen ”keskeinen juridinen sisältö” (2008, 152), joten kääntämisen kannalta tärkeää on, mitä missäkin tekstissä pidetään termin keskeisenä juridisena sisältönä ja riittävänä vastaavuutena. Esimerkiksi englannin kielen *marriage*-sanaa voi useimmissa tilanteissa pitää termin *avioliitto* lähivastineena.

Deskriptiivinen ekvivalentti tai *selittävä parafraasi* tarkoittaa termin käsitesisällön avaamista selittävin lisäyksin. Usein käännöksessä käytetään kohdeoikeusjärjestelmästä löytyvää laajempaa termiä ja sivulausetta, joka rajaa termin käsitesisällön vastaamaan lähdeoikeusjärjestelmän termiä (esimerkiksi termi *huoltaja* laissa lapsen huollosta ja tapaamisoikeudesta (1983/361) on ilmaistu englanninkielisessä käännöksessä ”person who has custody of a child”, ks. Act on Child Custody and Right of Access). Nimitystä *terminomainen selite* käytetään käännöksestä, jossa lähdekielinen termi korvataan lyhyellä selityksellä (Mattila 2008, 154). Esimerkiksi espanjankielinen termi *derecho foral* voidaan kääntää muotoon *Espanjan maakuntakohtaiset siviilioikeudelliset järjestelmät* (mp.).

Uudistermi eli *neologismi* on uusi omakielinen termi, jonka kääntäjä voi luoda silloin, kun kohdekielellä ei ole lähdekieliselle termille vastinetta. Uudistermi voi olla joko vierassana, käännöslaina tai uudissana (Mattila 2008, 150).

Luetteleva käännös on Mattilan käyttämä nimitys käännökselle, jossa lähdekielisen termin käsitesisältö saavutetaan kahdella tai useammalla kohdekielisellä termillä (Mattila 2008, 150). Mattila antaa esimerkiksi termin *säädös*, jolle ei ranskan kielellä ole vastinetta, mutta luetteleva käännös kuuluisi *lois, décrets et autres règlements* (mp.).

Sitaattilaina on termi, joka ilmaistaan lainatussa kirjoitusasussa joko sellaisenaan (esim. monilla EU-kielillä termi *acquis* lainataan sellaisenaan), yhdyssanan etuliitteenä (esim. ’franchising-toimilupa’) tai suluissa käännöksen perässä (esim. ’yhteisön säännöstö (*acquis*)’).

Eritoten viimeisellä keinolla lähinnä varmistetaan termin yksiselitteisyys, ja Mattilan mukaan esimerkiksi latinankielisiä sitaattilainoja käytettäessä termit olisi syytä myös kääntää, sillä myös lakimieslatinan termien käsitteissä on eri oikeuskulttuureissa huomattavia eroja (Mattila 2008, 151).

2.3.2 Oikeuskielen logiikan kääntäminen

Juridisten tekstien kääntäjiä on historiallisesti ohjeistettu kääntämään tekstejä sananmukaisesti (Šarčević 1997, 24). Krogsgaard Vesteragerin (2017, 107) mukaan juridisten tekstien ”pyhyys” on vaikuttanut niiden sananmukaisen kääntämisen suosioon paljon pidempään kuin muussa käännöstoiminnassa. Kinnunen sanoo juristien kuitenkin suhtautuvan edelleen varauksella käännöksiin sekä kääntäjien ammattitaitoon (2006, 33). Juridisten tekstien kääntäminen poikkeaa muista kääntämisen erikoisaloista myös sen vuoksi, että oikeuskielisiä tekstejä tulkitaan eri lähtökohdista kuin muiden alojen tekstejä. Šarčevićin mukaan (1997, 166) juristit ovatkin lähinnä kiinnostuneet syntaksin alla piilevistä loogisista päättelyketjuista eivätkä niinkään kieliopista sinänsä. Kinnunen listaa oikeusjärjestelmien ja juridisen ajattelutavan tuntemuksen osaksi myös juridisten tekstien kääntäjän kompetensseja (Kinnunen 2006, 134). Juridisten tekstien kääntäjän on kyettävä siis juridiseen ajattelutapaan tuottaakseen käännöksiä, joissa juridinen logiikka ja merkityssuhteet säilyvät samanlaisina lähdetekstistä kohdetekstiin.

Kääntäjältä vaadittavaan juridiseen ajattelutapaan sisältyy erityisesti sen tunteminen, miten oikeussäännöt muodostuvat. Oikeussääntöjen soveltamisessa juristit ja tuomarit käyttävät deduktiivista päättelytapaa, jota kutsutaan *sylogismiksi* (Sajama 2017, 47). Oikeudellisessa sylogismissa tehdään johtopäätös siitä, täyttääkö jokin tapaus oikeussäännössä esiintyvän *oikeusehdon* tai *tunnusmerkistön* (engl. *fact-situation*), jolloin normissa esitetty *oikeusseuraamus* tai *oikeusvaikutus* (engl. *statement of law*) pätee (Šarčević 1997, 137). Anne Lise Kjær esittää sylogistisen päättelyketjun yksinkertaisesti muodossa ”jos A on B ja B on C, A on C” ja oikeussäännön muodossa ”jos X niin Y” (Kjær 2000, 127).

Vaikka sylogismia on myös kritisoitu siitä, että se yksinkertaistaa tuomareiden järjestytyötä (ks. esim. Samuel 1994, 139), vaikuttaa sylogismi kuitenkin siihen, miten juridisten tekstien virkkeet ja merkityssuhteet rakentuvat. Oikeuskielessä, erityisesti *säädöskielessä*, käytetään tiettyjä kaavoja helpottamaan sylogistista deduktiota (Biel 2014a, 144). Esimerkiksi seuraavassa virkkeessä on sylogistisessa muodossa esitetty oikeusnormi:

If the default on payment is caused by the user's financial difficulties that he has run into because of serious illness, unemployment or some other special cause, principally through no fault of his own, the supply of electricity may be cut at the earliest two months after the due date of the payment.

Virkkeen aloittava sivulause määrittää oikeusehdon, jossa päälauseen osoittama oikeusseuraamus pätee. Bielin mukaan ehtolauseet, hypoteettiset muotoilut ja ehdolliset ilmaukset ovat tyypillisiä lainsäädännöllisissä teksteissä (mts. 144), ja yleisimmin päättelyketjut on ilmaistu samankaltaisina ”jos–niin” -ehtolauseina kuin edellä (mp.), vaikkakin on olemassa myös ehdottomia oikeussääntöjä, jotka eivät sisällä oikeusehtoa, vaan pelkän oikeusvaikutuksen (esim. monet perustuslailliset oikeusnormit, ks. Šarčević 1997, 147–148). *If*-konjunktion lisäksi muita ilmauksia, jotka indikoivat jos–niin-ehtolauseita, on englannin kielessä esimerkiksi *apart from, except, in so far as, in the absence of, in the case, in this event, notwithstanding, provided that, save for, unless, when, whenever, ja where* (Biel 2014a, 144).

Oikeussäännöt ovat useimmiten *preskriptioita* eli niissä ilmaistaan käsky, kielto tai lupa (Andström 2017). Šarčevićin mukaan oikeuskielissä teksteissä käytetään performatiivina usein sellaisia modaaliverbejä kuten *shall, shall not, may ja may not* eksplisiittisten performatiivien (esim. *order, forbid, permit*) sijaan (Šarčević 1997, 137). *Käsky* ilmaistaan englanniksi useimmiten modaaliverbiperformatiivilla *shall* (Šarčević 1997, 138). *Kielto* on käskyn vastakohta, joka ilmaistaan englanniksi performatiivilla *shall not* (Šarčević 1997, 146). *Lupa* kommunikoidaan performatiivilla *may* (mts. 142).

Käskystä Šarčević erottaa *edellytyksen*, joka taas ilmaistaan joko pääverbin preesensmuodon indikatiivilla tai performatiivilla *must* (1997, 138). Edellytys eroaa käskystä siten, että edellytys on luonteeltaan lähinnä prosessioikeudellinen ja että edellytyksen laiminlyönti vaikuttaa prosessin pätevyyteen siinä missä käskyn laiminlyönti on rangaistavaa (mp.). Edellytyksen ja käskyn erottaminen verbeillä *shall ja must* poikkeaa yleiskielestä, jossa verbit voidaan nähdä synonyymeinä. Edellytyksen ja käskyn performatiivien käyttö saattaakin täten poiketa huomattavasti englanninkielisissä lainsäädännön käännöksissä. Toisaalta myös Yhdistyneen kuningaskunnan parlamentin neuvonantajan toimisto (engl. *Office of the Parliamentary Counsel, OPC*) neuvoo nykyään välttämään *shall*-performatiivin käyttöä lainsäädännön laatimisessa ja suosittelee performatiivia *must* ilmaisemaan myös käskyä (*Drafting Guidance*, 14). Performatiivien käyttö ei ole siis täysin yksiselitteistä, ja eri instituutioissa käytännöt saattavat vaihdella. Juridisten tekstien käännöksissä vaihtelua performatiivien käytössä saattaa esiintyä vielä enemmän, ja myös esimerkiksi eksplisiittiset performatiivit, kuten *permit ja forbid*, saattavat olla yliedustettuina verrattuna alun perin kohdekielellä kirjoitettuihin

teksteihin, vaikka Šarčevićin mukaan niiden käyttöä vältetään lainsäädännöllisissä teksteissä (Šarčević 1997, 137).

2.4 Käännösuniversaalit

Käännösuniversaalit ovat ehdotettuja käännösten ”yleispiirteitä”, joita ilmenee lähde–kohdevertailussa sekä käännöksiä alun perin kohdekielellä kirjoitettuja teksteihin verrattaessa. Käännösuniversaalit voi Andrew Chestermanin (2010, 176–177) mukaan jakaa yleisesti kahteen ryhmään:

- 1) S-universaaleihin, jotka koskevat lähde- ja kohdetekstin eroja; ehdotettuihin S-universaaleihin kuuluvat muun muassa kohdetekstin pidentyminen lähdetekstiin verrattuna, kieltenvälinen interferenssi, standardisoituminen, eksplikoituminen ja toiston vähentyminen.
- 2) T-universaaleihin, jotka koskevat kohdetekstin tekstuaalisten piirteiden yhdenmukaisuutta alun perin kohdekielellä kirjoitettuihin teksteihin verrattuna; ehdotettuihin T-universaaleihin kuuluvat yksinkertaistuminen, normalisoituminen, kohdekielen epätyypilliset rakenteet sekä yksinomaan kohdekielelle ominaisten piirteiden aliedustuneisuus.

Käännösuniversaalien tutkimus on vauhdittunut erityisesti 1990-luvulta lähtien korpusmenetelmien yleistyttyä käännöstieteellisessä tutkimuksessa. Mona Bakerin mukaan vasta korpusmenetelmien myötä käännöstiede on saanut työvälineet käännöstekstien yleisten piirteiden tutkimiseen (Baker 1993, 243). Käännösuniversaaliteorian pätevyydestä kuitenkin myös kiistellään, minkä vuoksi on hyvä avata teoriasta käännöstieteellisessä kirjallisuudessa esitettyä kritiikkiä sekä pohtia käännösuniversaalihypoteesien soveltuvuutta vertailevan korpustutkimuksen apuvälineenä.

2.4.1 Käännösuniversaaliteorian kritiikkiä

Kääntämisen universaalien piirteiden luonne ja olemassaolo on kiistelty aihe, ja tutkimuksissa on saatu näyttöä sekä universaalihypoteesien puolesta että niitä vastaan (Mauranen 2004, 65). Useat tutkijat korostavat käännösuniversaalihypoteesien olevan nimenomaan hypoteeseja. Esimerkiksi Chesterman huomauttaa, että hypoteesien oletetaan pätevän kaikkiin käännöksiin kaikissa ajoissa, paikoissa ja kielissä, joten käännösuniversaaliteoriaa lienee mahdotonta todistaa (2004, 9). Sitäkin kyseenalaistetaan, johtuvatko käännösuniversaalit edes käännösprosessista itsestään. Juliane House pitää käännösuniversaaleja lähinnä kielen

universaaleina, joita esiintyy käännöksissä (2008, 11), ja on osoitettu, että käännösuniversaalien pätevyyden on osoitettu riippuvan kieliparista ja -suunnasta sekä tekstilajista (mts. 12). Myös Biel huomauttaa, että käännösuniversaaleihin tulisi suhtautua kriittisesti: käännösuniversaalihypoteesit ovat edelleen pelkkiä hypoteeseja, niihin liittyvät tutkimukset ovat olleet ristiriitaisia tai tuloksettomia, ja useimmat tutkimukset on tehty kaunokirjallisten tekstien käännöksistä, mikä ei kuvaa kaikkien kääntäjien, erityisesti juridisten tekstien kääntäjien, toimintaa (Biel 2014b, 183).

Erityisesti erikoisalojen teksteissä käännösuniversaalien pätevyys onkin kyseenalaistettu hanakammin verrattuna esimerkiksi kaunokirjallisten tekstien käännöksiin. Teoriaa ei myöskään ole määritelty kovin selkeästi, sillä esimerkiksi Chestermanin esittämät toiston vähentymisen ja yksinkertaistumisen hypoteesit ovat keskenään ristiriitaisia (toiston lisääntyminen kuuluu osaksi yksinkertaistumista, ks. Chesterman 2010, 177). Chestermanin mukaan (2004, 11) käännösuniversaalien tutkimuksessa tärkeintä ei liene sinänsä universaalius vaan käännösuniversaalihypoteesien avulla saatava tieto käännösten kielellisistä kaavoista. Käännösuniversaaleihin ei tulisikaan Maurasen (2006, 94) mukaan suhtautua absoluuttisina lakeina vaan lainomaisina tai todennäköisinä taipumuksina. Tällaisenaan käännösuniversaalihypoteesit voisivatkin toimia vertailevassa korpus tutkimuksessa apuna sen arvioimiseen, ovatko käännöskorpusten tekstuaaliset piirteet yhteensopivia alun perin kohdekielisten korpusten kanssa ja liittyvätkö tutkimuksessa löydetävät ilmiöt pikemmin käännösprosessin sisäisiin vai niistä riippumattomiin tekijöihin, kuten institutionaalsiin laatimisohjeistuksiin ja -käytäntöihin.

2.4.2 Vertaileva korpusmenetelmä ja T-universaalit

Kaikkia käännösuniversaalihypoteeseja ei kyetä tutkimaan vertailevalla korpusmenetelmällä. Vertailevassa korpus tutkimuksessa Chestermanin ehdottamista käännösuniversaaleista tavallisemmin tutkittu on T-universaalien ryhmä, sillä S-universaalien tutkiminen vaatisi lähde–kohde-vertailua, johon tarvitaan paralleelikorpusta. Vertailevalla korpusmenetelmällä tarkoituksena ei olekaan verrata kohdetekstien piirteitä lähdeteksteihin, vaan kohdetekstien tekstuaalista yhdenmukaisuutta alun perin kohdekielisiin teksteihin, minkä vuoksi menetelmä soveltuu paremmin Chestermanin luokittelemien T-universaalien tarkasteluun.

Toisaalta vaikka Chesterman on luokitellut S-universaaliksi eksplikoitumisen eli sen, että lähdetekstissä implisiittiset asiat saatetaan selittää käännöksessä eksplisiittisesti, liittyy eksplikoituminen juridisissa teksteissä myös oikeusjärjestelmien keskinäiseen terminologiseen

vastaamattomuuteen, jolloin se voitaisiin nähdä myös T-universaaliksi. Käsite, jolle ei löydy termiä kohdeoikeuskulttuurista, joudutaan selittämään auki eksplisiittisesti, kun taas lähdeoikeuskulttuurin termi on implisiittinen lähdetekstin lukijalle, joka tuntee lähdeoikeuskulttuurin käsitejärjestelmän.

Täten esimerkiksi suomalaista oikeusjärjestelmää kuvaavat englanninkieliset tekstit ovat todennäköisesti eksplisiittisempiä kuin alun perin englanniksi kirjoitetut *common law* -oikeusjärjestelmää kuvaavat tekstit. Myös Biel on sitä mieltä, että eksplikoituminen voi vaikuttaa myös T-universaalin tavoin käännöksen tekstuaalisten piirteiden yhteensopivuuteen kohdekielisten tekstien joukossa (Biel 2014a, 101). Tiina Puurtisen (2004, 166) mukaan käännöksissä eksplikoituminen näkyy suurempana määränä lauseen yhdistäviä lauseenjäseniä, kuten konjunktioita, tiettyjä adverbeja sekä relatiivipronomineja. Näiden lisäksi juridisten käännösten tekstuaalisia piirteitä tarkasteltaessa eksplikoituminen saattaa näyttäytyä käännöksissä alun perin kohdekielellä kirjoitettuihin teksteihin verrattuna myös pienempänä leksikaalisena tiheytenä eli sisältösanojen suhteena kaikkiin sanoihin sekä suurempana tarkenteiden käyttönä.

Lisäksi vertailevan korpustutkimuksen kannalta kiinnostavia T-universaaleja ovat erityisesti *yksinkertaistuminen* ja *normalisoituminen*, sillä niitä kyetään mittaamaan korpuksista helposti. *Yksinkertaistuminen* tarkoittaa, että käännöksillä on taipumus olla alun perin kohdekielellä kirjoitettuja tekstejä yksinkertaisempia. Tämä näyttäytyy Chestermanin mukaan pienempänä TTR-lukuna, pienempänä leksikaalisena tiheytenä, ja suurempana osuutena suuren frekvenssin yksiköitä (Chesterman 2010, 177). *Normalisoituminen* liittyy Bielin (2024, 103) mukaan siihen, että käännöksissä tekstuaalisten piirteiden vaihtelu on pienempää esimerkiksi virkepituuksien ja TTR-lukujen osalta, ja alun perin kohdekielisissä teksteissä vaihtelua esiintyy enemmän kuin käännöksissä (mp.).

Vertailevassa korpustutkimuksessa T-universaaleja voidaan käyttää apuna sen mittaamiseen, onko käännösteksteistä koostuva korpus tekstuaalisilta piirteiltään yhdenmukainen alun perin kohdekielisistä teksteistä koostuvien korpusten kanssa, vai ovatko käännökset esimerkiksi yksinkertaistuneita tai normalisoituneita alun perin kohdekielisiin teksteihin verrattuna. Tällöin tarkastelun keskiössä eivät ole T-universaalit sinänsä, vaan hypoteeseja käytetään suuntaviivoina sille, miten käännösten tekstuaalisten piirteiden voisi olettaa suhteutuvan alun perin kohdekielisiin teksteihin. Mikäli käännöskorpus näyttäytyy selvästi käännösuniversaalihypoteeseista poikkeavana, saattavat erot johtua esimerkiksi tekstilajista,

institutionaalisista tekstin laatimisen käytännöistä tai muista käännösprosessin ulkopuolisista tekijöistä.

3 Menetelmä ja aineisto

Tässä luvussa kuvailen tutkimuksessa tarkasteltavan aineiston sekä käytetyn tutkimusmenetelmän. Tutkielman aineistona on kolme korpusa, joihin on kerätty säädöksiä seuraavista lähteistä (lista aineistoon valituista säädöksistä liitteessä 1):

- 1) suomalaisten *lakien* englanninkielisiä käännöksiä sähköisestä Finlex-säädöskokoelmasta (myöhemmin *FI-korpus*)
- 2) Euroopan unionin *asetusten* englanninkielisiä toisintoja EUR-Lex-oikeusportaalista (myöhemmin *EU-korpus*) sekä
- 3) Yhdistyneen kuningaskunnan lakeja (*Acts*) Legislation.gov.uk-palvelusta (myöhemmin *UK-korpus*).

Taulukkoon 1 alla on listattu korpusien sanamäärät sekä yksittäisten tekstien lukumäärä kussakin korpuksessa.

Taulukko 1. Korpusien sanamäärät ja korpustekstien lukumäärät

	FI-korpus	UK-korpus	EU-korpus
sanamäärä	198 955	196 368	196 909
tekstien lukumäärä	26	18	24

Kuten aiemmin on mainittu, suomalaisten lakien englanninkieliset käännökset ovat epävirallisia, mikä tarkoittaa sitä, että ne eivät ole juridisesti sitovia tai todistusvoimaisia samoin kuin suomen- ja ruotsinkieliset toisinnot suomalaisista laeista. Yhdistyneen kuningaskunnan lait taas kirjoitetaan alun perin englanniksi ja myös Euroopan unionin asetukset useimmiten englanti alkuperäisenä kielenä. Näiden kahden muun korpuksen tekstit ovat siis virallista, juridisesti sitovaa lainsäädäntöä, mikä mutkistaa tutkimusasetelmaa; aineistossa on eri periaatteilla kirjoitettuja tekstejä, jotka epäilemättä kuuluvat samaan tekstilajiin mutta joiden muodostamisperiaatteet eroavat toisistaan. Säädösten epäviralliset käännökset ovat funktioltaan deskriptiivisiä, kun taas juridisesti sitovat, auktorisoidut säädöskäännökset ovat preskriptiivisiä (Šarčević 1997, 11). Šarčevićin mukaan säädösten epävirallisia käännöksiäkin käytetään ulkomaisissa tuomioistuimissa, joten niilläkin on juridista painoarvoa (mts. 277), ja niiden kääntämisessä tulisi soveltaa samoja periaatteita kuin usealla kielellä lainvoimaisen lainsäädännön kääntämisessä (mts. 279). On kuitenkin huomioitava, että suomalaisten lakien epävirallinen käännöstyö ministeriöissä poikennee

huomattavasti esimerkiksi siitä, miten ruotsinkieliset käännökset tuotetaan suomalaisissa lainvalmisteluelimissä. Eräänä mielenkiinnon kohteena tutkimuksessa onkin se, näkyykö kääntäjän ”eristyneisyys” lainvalmisteluprosessista epävirallisissa käännöksissä verrattuna EU-asetuksiin ja alun perin englanniksi laadittuihin lakeihin.

Aineisto on rajattu niin, että tekstejä on koottu vuonna 1994 ja sitä myöhemmin annetuista ympäristöoikeuden ja energiaoikeuden alojen säädöksistä. Oikeusalarajoituksen tarkoituksena on varmistaa korpusten vertailtavuus, ja ympäristö- ja energiaoikeudellisia tekstejä löytyy helposti kustakin tämän tutkimuksen oikeusjärjestelmästä. Tutkimuksen tarkoituksena on tarkastella Euroopan unioniin liittymisen jälkeistä juridista käännösonglantia Suomessa ja verrata sitä lainvoimaiseen englanninkieliseen lainsäädäntöön. Vaikka Suomi liittyi Euroopan unioniin vuonna 1995, päätin ottaa mukaan kaksi suomalaista lakia myös vuodelta 1994 aineiston laajentamiseksi. Käytännössä tällä ei pitäisi olla merkitystä, sillä valmistelut Suomen liittymiseksi Euroopan unioniin ovat olleet jo silloin vauhdissa (esimerkiksi käännöstyö EY:n säännöstöä varten on aloitettu jo 1990-luvun alussa, joskin ranskasta englannin sijaan, ks. Vuorinen 2017, 11–12). Kaikkiin Finlex-säädöskokoelman käännöksiin ei ole myöskään dokumentoitu julkaisupäivämäärää, joten vuonna 1994 annetun lain käännös saattaa silti olla julkaistu vuonna 1995 tai sitä myöhemmin. Tässä tutkimuksessa en ole myöskään kiinnostunut ajallisesta pitkittäisvertailusta, osaksi juuri siitä syystä, että suomalaisten lakien käännöksiä ei ole päivätty tarkasti eikä niihin tehtyjä muutoksia ole dokumentoitu riittävästi.

En ole rajannut aineistoon valittavia säädöksiä pituuden mukaan, mutta olen pyrkinyt muodostamaan korpukset niin, että tutkimuksen korpuksissa sanamäärät olisivat mahdollisimman samankaltaisia. Pidemmät tekstit yhdessä korpuksessa muihin nähden saattavat näyttäytyä pienempänä TTR-lukuna eli yksittäisten sanojen suhteena kaikkien saneiden lukumäärään (engl. *type-token ratio*, *TTR*), mikäli samat sanat toistuvat yksittäisessä pitkässä säädöksessä usein. Muuten pidempien korpustekstien ei pitäisi olennaisesti vaikuttaa tutkimuksen tuloksiin.

3.1 Tutkimuskysymys

Tutkimuksen tarkoituksena on muodostaa käsitys siitä, onko suomalaisten säädösten epävirallisten käännösten englanti olennaisesti erilaista verrattuna Euroopan unionin ja Yhdistyneen kuningaskunnan säädösenglantiin. Tekstuaalisten piirteiden vertailun helpottamiseksi havaintoja peilataan *eksplikoitumisen*, *normalisoitumisen* ja *yksinkertaistumisen* käännösuniversaalihypoteeseihin.

Hypoteesina on, että Suomen kuuluessa mannereurooppalaiseen oikeusryhmittymään käytettävä säädösenglanti muistuttaa pikemmin Euroopan unionin lainsäädännön laatimiseen käytettyä englantia, sillä vaikka suurin osa Euroopan unionissa laadittavista teksteistä laaditaan alun perin englanniksi, laativat tekstejä suurelta osin jotakin muuta kieltä kuin englantia äidinkielenään puhuvat henkilöt, minkä vuoksi EU-tekstit saattavat jossakin määrin vaikuttaa tekstuaalisilta piirteiltään käännöksiltä. Mikäli käännösuniversaalit pätevät juridisten tekstien käännöksiin samoin kuin esimerkiksi kaunokirjallisuuteen, pitäisi suomalaisten englanninkielisten käännösten kuitenkin esiintyä myös EU-teksteihin verrattuna selkeästi eksplikoituneempina, normalisoituneempina ja yksinkertaistuneempina.

Hypoteesia testataan kahdella tutkimuskysymyksellä:

- 1) Miten käännetty kansallinen lainsäädäntö eroaa kohdekielisestä ylikansallisesta lainsäädännöstä ja kohdekielisestä kansallisesta lainsäädännöstä:
 - a. makrotasolla (virkepituus, sanaluokat, leksikaalinen tiheys, TTR-luku)?
 - b. mikrotasolla (sanasto, n-grammit, juridisen logiikan kommunikointi)?
- 2) Pätevätkö *eksplikoitumisen*, *normalisoitumisen* ja *yksinkertaistumisen* käännösuniversaalihypoteesit lainsäädännön käännöksissä?

3.2 Korpusmenetelmä ja korpustyökalut

Korpusmenetelmiä on käytetty käännöstieteellisen tutkimuksen osana kasvavasti 1990-luvulta alkaen (Biel 2014a, 95). Biel määrittelee korpuksen ”koneellista kielitieteellistä analyysia varten kootuksi edustavaksi kokoelmaksi luonnollisesti esiintyvää kieltä” (mp. [oma käännös]). Korpusmenetelmät tarjoavatkin kätevän tavan hankkia empiiristä, kvantitatiivista tietoa kielestä. Korpuksen pohjalta tehtävät yleistyksen ja johtopäätökset voivat kuitenkin koskea ainoastaan tutkittua otosta, ei koko kieltä (Biel 2014a, 89). Korpusten avulla voidaan kuitenkin saada selville toistuvia ilmiöitä ja kielellisiä kaavoja, joiden manuaalinen etsiminen olisi

työlästä, vaikkakin pelkkä korpustutkimus ei korvaa kvalitatiivista käännösten tutkimusta. Bhatia (2004, 224) kiteyttää korpusanalyysin rajoitteet toteamalla, että kvalitatiivinen analyysi alkaa siitä, mihin korpuslingvistiikka jää. Tässä tutkimuksessa en yhdistä laadullisia menetelmiä korpusanalyysiin, mutta pyrin selittämään saatuja tutkimustuloksia vertailevan oikeuslingvistiikan ja juridisen kääntämisen tutkimuksen näkökulmista. Korpusten tekstuaalisten piirteiden erojen tarkastelussa käytän apuna myös käännösuniversaalioteoriaa (ks. luku 2.4). Käännösuniversaalioteorian ja eri käännösuniversaaliyhypoteesien pätevyys on kiistanalaista, mutta käännösuniversaaliyhypoteesit toimivat hyvin eräänlaisena komparatiivisten korpusten mittatikkuna: niiden avulla voidaan tehdä johtopäätöksiä siitä, johtuvatko mahdolliset erot korpusten tekstuaalisissa piirteissä käännösprosessin sisäisistä vai ulkoisista piirteistä. Toisin sanoen, jos korpustutkimuksen tulokset ovat olennaisesti ristiriidassa käännösuniversaaliyhypoteesien kanssa, syitä eroavaisuuksiin löytynee todennäköisesti käännösprosessin ulkoisista tekijöistä, kuten tekstien institutionaalisten laatimiskäytäntöjen tai oikeusjärjestelmien eroista.

Korpusanalyysiin olen käyttänyt *Sketch Engine* -korpustyökalua, joka on internetselaimella toimiva korpuspalvelu, sekä apuna osittain myös *Wordsmith*-korpustyökalua. Sketch Enginellä on mahdollista ajaa yleisimpiä korpusanalyyseja, kuten konkordanssihakua sekä n-grammien ja sanojen frekvenssilistoja. Sketch Enginellä ei ole mahdollista ajaa analyyseja liittyen standardisoituun TTR-lukuun ja virkepituuksiin, joten näihin analyysihin on käytetty apuna Wordsmith-ohjelmaa. Tässä tutkimuksessa aion tarkastella korpuksista seuraavia tekstuaalisia piirteitä:

- 1) virkepituus
- 2) sanaluokkajako
- 3) leksikaalinen tiheys
- 4) TTR-luku
- 5) sanalistat
- 6) juridista logiikkaa kommunikoivat ilmaukset sekä
- 7) n-grammit.

Korpusten virkkeiden jäsentymistä tarkastellaan vertailemalla keskimääräisiä virkepituuksia sekä virkepituuden keskihajontaa. Oikeuskielen on ehdotettu olevan muita erikoiskieliä pitkävirkkeisempää (Mattila 2017, 160), joten myös virkkeiden pidentymistä tai lyhentymistä käännöskorpuksissa pitäisi olla helppo havainnoida. Juridisten tekstien käännöksissä käytetään

lisäksi usein selittäviä parafraaseja ja eksplikointia eri oikeuskielten ja -järjestelmien terminologisesta vastaamattomuudesta johtuen (Šarčević 1997, 150–151; Biel 2014a, 100). Vaikka oletan tämän osaltaan pidentävän virkkeitä käännöksissä, saattavat käännökset olla yksinkertaistumishypoteesin mukaan myös lyhyempivirkkeisiä (Biel 2014a, 101), joten virkepituus ei välttämättä anna yksiselitteistä tietoa eksplikoitumisesta tai yksinkertaistumisesta. Mahdolliset erot virkepituuksien vaihtelussa saattavat kuitenkin kieliä käännösten normalisoitumisesta.

Sketch Enginen tekemä sanaluokka-annotointi mahdollistaa korpusten rakenteiden vertailun sanaluokkatasolla sekä *leksikaalisen tiheyden* laskemisen korpuksista. Sanaluokkien tarkastelussa keskeistä on eksplikoitumishypoteesin näkökulmasta konjunktoiden, substantiivien, pronomien ja prepositioiden käyttö. Oletuksena on, että lainsäädännön epävirallisissa käännöksissä käytetään konjunktioita, prepositioita ja pronomineja termejä selittävässä parafraaseissa, mikä tekee käännöksistä eksplisiittisempiä ja myös pienentää leksikaalista tiheyttä. *Leksikaalinen tiheys* tarkoittaa leksikaalisten sanojen – toisin sanoen substantiivien, verbien, adjektiivien ja sisällöllisten adverbien – suhdetta kaikkiin teksteissä esiintyviin sanoihin. Yksinkertaistumishypoteesin mukaan käännöksissä leksikaalisen tiheyden on osoitettu olevan pienempi kuin alun perin kohdekielellä kirjoitetuissa teksteissä (Laviosa 2003, 153). Ojalan (2015, 49) tutkimuksessa havaittiin kuitenkin Laviosan tutkimustulosten kanssa ristiriitaisesti, että leksikaalinen tiheys on korkeampi Euroopan unionin asetusten suomenkielisissä toisinnoissa kuin alun perin suomeksi kirjoitetuissa suomalaisissa säädöksissä, joten juridiset tekstit eivät välttämättä näyttäydy leksikaaliselta tiheydeltään yksinkertaistumis- ja eksplikoitumishypoteesin mukaisiksi.

TTR-luku eli *sana–sane-suhde* (engl. *type-token ratio*) tarkoittaa yksittäisten sananmuotojen osuutta kaikista korpuksen saneista. Yksinkertaistumishypoteesin mukaan toistosta johtuen käännösten TTR-luvun tulisi olla alun perin kohdekielisiä tekstejä pienempi, toisin sanoen yksinkertaistuneessa käännöksessä pienempi osa sananmuodoista muodostaa kaikki tekstin saneet. TTR-luvun rinnalle lasketaan tuhannen sanan jaksoissa standardisoitu TTR-luku (STTR), joka antaa korpusten sana–sane-suhteesta paremman, yksittäisten tekstien pituudesta riippumattoman kokonaiskuvan. Standardisoidun TTR-luvun keskihajonnan avulla kyetään myös tarkastelemaan, ovatko käännökset TTR-luvun suhteen normalisoitumishypoteesin mukaisia.

Sanalistat ovat korpuksista ajettuja sananmuotojen frekvenssilistoja, joista verrataan funktiosanojen käytön eroja korpusten välillä. Yksinkertaistumishypoteesin mukaan käänöksissä tulisi olla suurempi osuus suuren frekvenssin yksiköitä, ja sanalistaista pystytään myös tarkastelemaan eksplikoitumiseen liittyvien prepositioiden, konjunktoiden ja pronomien käyttöä.

Sanojen n-grammeilla viitataan siihen, että tietyt sanat muodostavat toistuvia n yksikön mittaisia jonoja korpuksessa. Koska fraseologismien – jotka ovat oikeuskielessä erityisen yleisiä (Biel 2014a, 34–36; Mattila 2017, 157) – voi olettaa näkyvän n -grammianalyysissä, on mielekästä tutkia korpusten välisiä eroja n -grammien frekvensseissä. N -grammianalyysin avulla saadaan todennäköisesti tietoa myös siitä, millä eri tavoilla lainsäädännöllisiä tekstejä jäsennetään eri instituutioissa. N -grammien analyysissä täytyykin ottaa huomioon erilaiset laatimiskäytännöt ja kaavamaiset ilmaisut, jotka saattavat näkyä tuloksissa huomattavan paljon. Yksinkertaistumishypoteesiin perustuen käänöskorpuksessa suurifrekvenssisiä n -grammeja tulisi olla enemmän kuin alun perin kohdekielisissä korpuksissa.

Juridisen logiikan kommunikoinnin analyysin tarkoituksena on löytää juridisiin ehtolauseisiin sekä velvollisuuksien, kiellon ja luvan kommunikointiin liittyviä sanoja ja ilmauksia, jotka ovat yli- tai aliedustuneita käänöskorpuksessa kahteen muuhun korpukseen verrattuna. Tarkasteltavien juridista logiikkaa kommunikoivien ilmausten analysointiin käytetään konkordanssihakua. Yksinkertaistumishypoteesin mukaan myös juridisen logiikan kommunikointiin liittyvän sanaston tulisi olla yksinkertaistunutta käänöksissä.

Analyyseista saatujen tulosten tilastollista merkitsevyyttä on arvioitu kahdella eri testillä: keskiarvojen erojen merkitsevyyttä virkepituudessa ja TTR-luvussa on arvioitu t -testillä ja sanaston, n -grammien, leksikaalisen tiheyden sekä sanaluokkien eroja χ^2 -testillä käyttäen kaksi kertaa kahden ruudun matriisia. Testeissä nollahypoteesina on, että erot eri korpuksista saaduissa arvoissa ovat satunnaisten vaihtelun tulosta. Vastahypoteesina sen sijaan on, että erot korpusten välisissä arvoissa eivät johdu satunnaisesta vaihtelusta. Nollahypoteesille on laskettu P -arvo, joka ilmaisee todennäköisyyttä, jolla nollahypoteesi pätee. Tässä tutkimuksessa käytetään seuraavia tilastollisen merkitsevyyden raja-arvoja:

- P -arvon ollessa alle 0,05 on ero tilastollisesti melkein merkitsevä.
- P -arvon ollessa alle 0,01 on ero tilastollisesti merkitsevä.
- P -arvon ollessa alle 0,001 on ero tilastollisesti erittäin merkitsevä.

Korpusten suuresta koosta johtuen P-arvot osoittavat suuressa osassa korpusten välisiä vertailuja erittäin suurta tilastollista merkitsevyyttä, mistä johtuen erityisen mielenkiintoista tässä analyysissa on tarkkailla suuria P-arvoja, jotka ilmaisevat todennäköistä samankaltaisuutta korpusten välillä tarkasteltujen ilmiöiden suhteen. Näin tilastollista merkitsevyyttä kyetään käyttämään apuna tutkimustulosten merkittävyyden arvioinnissa. Toisin sanoen, tutkimuksen tuloksista tilastollisesti merkitsemättömät erot ovat vertailun kannalta merkittäviä.

3.3 Säädösten rakenteet

Tutkimuksessa verrataan kolmen eri oikeusjärjestelmän säädöksiä, jotka eroavat rakenteeltaan toisistaan huomattavasti. Tämän vuoksi olen mukaillut Aino Piehlin korpustutkimusta suomenkielisten EU-asetusten ja suomalaisten lakien eroista (2006) ja sisällyttänyt aineistoon valituista säädöksistä vain varsinaisia säännöksiä. Teksteistä on täten poistettu korpukseen syöttämisen yhteydessä säädösten liitteitä ja muita osia, jotka saattaisivat vääristää tutkimuksen tuloksia. Tällaisia osia ovat Yhdistyneen kuningaskunnan lakien alussa oleva täydellinen otsikko (*long title*), joka esittelee lain tavoitteet ja perustelut yhdessä virkkeessä, kaavamainen ilmaus ”BE IT ENACTED by the Queen’s most Excellent Majesty, by and with the advice and consent of the Lords Spiritual and Temporal, and Commons, in this present Parliament assembled, and by the authority of the same, as follows - -” sekä Euroopan unionin asetusten alussa oleva johdanto-osa, jossa esitellään säädöksen tavoitteet ja perustelut säädöksen antamiselle sekä säädöksen antava elin. Nämä osat periaatteessa kietovat säädöksen yhtenäiseksi *makrovirkkeeksi*. Suomalaisten säädösten englanninkielisistä käännöksistä on poistettu huomautus käännöksen epävirallisuudesta ja käännöksen toimeenpanneen ministeriön tiedot. Kaikista aineiston teksteistä on poistettu myös säädöksen nimi ja liitteet mukaan lukien Yhdistyneen kuningaskunnan lakeihin liitetyt *schedules*-osiot, joissa kuvataan tarkemmin, miten lain säännösten on tarkoitus toimia käytännössä (*Schedules*).

Ojala (2015) on tutkimuksessaan poistanut korpusteksteistä myös voimaantulosäännökset, sillä ne toistuvat usein kaavamaisena esimerkiksi Euroopan unionin asetuksissa, mutta tässä tutkimuksessa olen jättänyt ne aineistoon. Tähän on kolme syytä. Ensiksi, kaavamaisena toistuva osa Euroopan unionin asetuksissa on usein vain lyhyehkö lause, joka ei yksistään juuri vääristä tuloksia: ”This Regulation shall enter into force on the twentieth day following that of its publication in the Official Journal of the European Union”. Kaikissa asetuksissa voimaantuloa ei ole myöskään annettu tässä muodossa, vaan esimerkiksi ilmaisemalla tarkka

päivämäärä. Voimaantulosäännös saattaa näkyä korpuksen n-grammianalyysissä yksittäisenä sanajonona, joka on helppo erottaa frekvenssilistasta ja ottaa tulosten tarkastelussa huomioon. Toiseksi, muissa osa-aineistoissa voimaantulosäännökset eivät ole samalla tavalla kaavamaisia. Yhdistyneen kuningaskunnan laeissa usein voimaantulosäännökseen on lisätty esimerkiksi alueet, joita laki koskee, joten olisi haasteellista määrittää, mitä osia voimaantulosäännöksistä karsittaisiin pois kunkin korpuksen kohdalla. Kolmanneksi, kaikkien korpusten teksteissä on myös varmasti muita kaavamaisina toistuvia ilmauksia, joita olisi työlästä etsiä ja erottaa aineistosta, ja myös kaavamaisien ilmausten voi perustella kuuluvan olennaiseksi osaksi kielenkäyttöä oikeuskonteksteissa. Esimerkiksi Mattila (2017, 117) pitää kaavamaisuutta, joka ilmenee esimerkiksi valmiiden tekstipohjien hyödyntämisenä, yhtenä oikeuskielen universaaleista piirteistä. Lisäksi haluan tutkimuksessani kohdella aineiston eri osia mahdollisimman johdonmukaisesti, joten voimaantulosäännökset on jätetty kaikkiin osa-aineistoihin sellaisenaan. Euroopan unionin asetuksista olen kuitenkin poistanut asetuksen allekirjoittaneet tahot ja allekirjoituksen päivämäärät. Myös suomalaisten lakien käännöksistä niissä mahdollisesti esiintyvät allekirjoittajien nimet on poistettu.

3.4 Suomalaiset lait

FI-korpus koostuu 26 eri lain englanninkielisestä käännöksestä, ja olen aineistonkeruussa hyödyntänyt Finlex-palvelun tarkennettua hakua sekä englanninkielisten käännösten asiasanaindeksointia. Haku on tehty hakutermeillä *energy*, *power*, *environment**, *forest**, *waste*, *electric**, ja *climate*; aineistoa on täydennetty asiasanojen *waste*, *environment*, *environmental impact*, *forestry*, ja *conservation of nature* alle luokitelluilla käännöksillä. Olen sisällyttänyt suomalaisten lakien korpukseen kaikki näillä asia- ja hakusanoilla löytyvät englanninkieliset käännökset vuodesta 1994 eteenpäin, joten suomalaisten lakien korpuksen voi katsoa muodostavan eräänlaisen kokonaisaineiston. Muut korpuksien muodostettu satunnaisotantana, sillä hakutulokset EUR-Lex- ja Legislation.gov.uk-palvelussa ovat laajempia. Finlex-palvelun puutteellinen asiasanoitus ja ainoastaan säädösten otsikoista hakeva hakuliittymä rajoittaa korpuksen kokoa. On otettava huomioon myös, että kaikkia säädöksiä ei ole käännetty englanniksi, tai ne eivät ainakaan ole saatavilla Finlex-palvelun kautta. Sädösten sisällöistä voisi löytää hakutuloksia esimerkiksi Google-hakukoneella, mutta en ole sisällyttänyt aineistooni tällaista hakua, sillä aineisto on näinkin riittävä. Piehlin tutkimuksen (2006) korpuksessa on yhteensä 124 361 juoksevaa sanaa eli *sanetta* suomalaisia lakeja ja suomenkielisiä EU-asetuksia. On tietysti huomioitava, että suomi on synteettinen kieli, ja suomenkielisistä teksteistä muodostetut korpuksien ovat vääjäämättä sanamääräisesti katsottuna

pienempiä. Tässä suomalaisten lakien englanninkielisten käännösten korpuksessa on 198 955 juoksevaa sanaa, joka on enemmän kuin Piehlin alakorpuksissa yhteensä. Kokonaisaineisto antaa nimensä mukaisesti kokonaiskuvan tarkasteltavasta perusjoukosta, mutta ottaen huomioon, että kyseessä on erikoiskielen korpus, voi siitä tehdä myös varovaisia yleistyksiä suomalaisten säädösten englanninkielisistä käännöksistä yleisemminkin. Tämän tutkimuksen muut korpuksat (Yhdistyneen kuningaskunnan lait ja Euroopan unionin asetukset) on koottu yksinkertaisena satunnaisotantana vastaamaan sanamääriltään suomalaisten lakien korpusta.

3.5 Yhdistyneen kuningaskunnan lait

Yhdistyneen kuningaskunnan lait olen kerännyt UK-korpukseen Legislation.gov.uk-lainsäädäntöportaalista. Palvelun hakuliittymä on monipuolisempi kuin Finlex-palvelussa, ja palvelun kautta on saatavilla kaikki säädökset vuodesta 1988 eteenpäin. Haun voi rajata vuoden ja säädöstyypin mukaan, ja hakusanoilla voi etsiä joko säädösten otsikoista tai sisällöstä. Yhdistyneen kuningaskunnan lakien poiminnassa käytettiin samoja hakusanoja kuin suomalaisten säädösten hakemiseen, ja säädöstyypin rajattiin koskemaan Yhdistyneen kuningaskunnan yleisiä lakeja (*UK Public General Acts*) vuodesta 1994 alkaen. Aineiston ulkopuolelle rajautuu siis niin kutsutusti paikalliset lait (*UK Local Acts*), joita asetetaan usein mahdollistamaan yhtiöiden muodostamista tai rakennusprojekteja (esim. *New Southgate Cemetery Act 2017* ja *Bank of Ireland (UK) plc Act 2012*).

Palvelussa ei ole listattu säädöksiä asiasanoittain, joten tein haun samoilla hakusanoilla kuin suomalaisten lakien kohdalla. Näillä hakusanoilla löytyi 235 eri lakia, sillä otsikkohaku hakee myös pitkistä otsikoista. Tein näiden hakutulosten pohjalta ensin listauksen aineistoon oikeusalan puolesta soveltuvista teksteistä (yhteensä 76 lakia), jonka joukosta valitsin yksinkertaisena satunnaisotantana 26 lakia. Koska monet laeista ovat aihealueiltaan yleisempiä kuin suomalaiset lait (monesti tämän voi päätellä nimestä, esim. *Energy Act 2013*) ja täten huomattavasti pidempiä, jouduin poistamaan otoksesta kahdeksan tekstiä, jotta korpus olisi sanamäärältään vastaava muiden korpusten kanssa.

UK-korpuksessa on 18 tekstiä ja yhteensä 196 368 juoksevaa sanaa. Lukumääräisesti UK-korpuksessa on vähemmän tekstejä, koska osa teksteistä on huomattavasti pidempiä kuin mikään teksti muissa korpuksissa. Halusin kuitenkin jättää korpukseen aihepiiriltään korpuksen kannalta olennaisia tekstejä, kuten *Environment Act 1995*, vaikkakin kyseisessä laissa yksistään on yli 60 tuhatta juoksevaa sanaa. Toinen laki korpustekstien pidemmästä päästä on *Energy Act*

2013, jossa on reilut 45 tuhatta juoksevaa sanaa. Nämä kaksi lakia siis muodostavat suurimman osan korpuksesta, mutta koska lakien sisältö on laaja ja oikeusalarajauksen kannalta olennainen, ei niiden pois pudottaminen olisi ollut perusteltua.

3.6 Euroopan unionin asetukset

EUR-Lex-oikeusportaalista asetukset on poimittu niin ikään satunnaisotantana. Hakuliittymästä pystyy selaamaan asetuksia aihepiirien alta, joten rajasin haun vuodesta 1994 alkaen annettuihin asetuksiin asiasanojen *ympäristö* ja *energia* alta. En ole rajannut asetuksia sen mukaan, onko ne antanut Euroopan unionin neuvosto, Euroopan parlamentti vai Euroopan komissio. Näillä kriteereillä löytyi yhteensä 159 asetusta, joista korpukseen poimittiin yksinkertaisena satunnaisotantana 26 kappaletta, mutta kaksi asetusta pudotettiin otoksesta pois, jotta korpus saatiin sanamäärältään vastaamaan suomalaisten lakien korpusta. Euroopan unionin asetusten korpuksessa on yhteensä siis 24 tekstiä ja 196 909 juoksevaa sanaa.

Korpuksat ovat siis sanamääräisesti katsottuna pitkälti toisiaan vastaavia. Yhdistyneen kuningaskunnan lakien korpus on sanamääräisesti 0,12 prosenttia suurempi kuin suomalaisten lakien korpus, ja Euroopan unionin asetusten korpus taas on 0,11 prosenttia sitä pienempi. Yhdistyneen kuningaskunnan lakien korpuksessa on kuitenkin pienempi määrä yksittäisiä tekstejä, minkä ei kuitenkaan pitäisi vaikuttaa tässä tutkimuksessa tarkasteltaviin korpusten ominaisuuksiin.

3.7 Aineiston syöttäminen korpukseen

Aineiston tekstit on syötetty kolmeen eri korpukseen TXT-tiedostoina merkistökoodauksella UTF-8. Joillakin asetuksilla korpus- ja annotointityökalut tulkitsevat rivinvaihdon virkkeen päättymiseksi, mikä olisi osoittautunut ongelmalliseksi tutkimuksen kannalta, sillä lainsäädännössä yleisesti yksittäiset säännökset on aseteltu usealle riville, erityisesti Yhdistyneen kuningaskunnan laeissa. Tässä tutkimuksessa on käytetty Sketch Engineen oletuksena sisältyvää sanaluokka-annotointia (*English 3.3 for TreeTagger pipeline v2*), jossa ainoastaan virkkeen päättävä välimerkki, toisin sanoen piste, huutomerkki ja kysymysmerkki, merkitsee uuden virkkeen alkamisen. Tämän ansiosta pitkät, useisiin alakohtiin jaetut oikeuskieliset virkkeet annotoituvat korpuksessa oikein, eikä näin ollen ole tarvetta poistaa rivinvaihtoja korpusteksteistä. Myös Wordsmith-analyysit ajettiin samoille Treetaggerilla annotoiduille korpusteksteille.

Aineiston teksteissä on paljon ristiviittauksia muihin lakeihin numeerisessa muodossa (esim. 109/2003), jotka korpustyökalu tulkitsee erillisiksi sanoiksi. Tämä on hyvä ottaa huomioon tutkimustuloksia analysoitaessa. Ojala (2015, 35) on todennut, että kyseessä on lakikielelle tyypillinen ilmiö, jonka ei pitäisi huomattavasti vaikuttaa kielellisten piirteiden vertailuun. Huomattavin vaikutus tällä piirteellä on todennäköisesti virkepituuksiin, leksikaaliseen tiheyteen sekä numeraalien lukumäärään sanaluokkakajakaumissa. Toinen huomion arvoinen seikka korpusten rakenteessa on, että pykälien, artiklojen ja osioiden otsikot luetaan korpuksessa otsikkoa seuraavaan virkkeeseen kuuluviksi. Otsikot saattavat marginaalisesti vaikuttaa korpusten virkepituuksiin, mutta koska tilanne on kaikissa korpuksissa sama, ei otsikoiden pitäisi olennaisesti vaikuttaa korpusten vertailtavuuteen. Ottaen nämä seikat huomioon otsikoiden erottaminen virkkeistä manuaalisesti olisi työlästä eikä kovin tarkoituksenmukaista.

Korpuksissa on siis tiettyjä rajoitteita, jotka täytyy huomioida tutkimustuloksia tarkasteltaessa. Korpustutkimukseen liittyy aina aineiston edustavuuden määrittämisen ongelma, eikä manuaalinen sanaluokkien ja virkkeiden annotointi ole käytännössä mahdollista, vaikka tämän tutkimuksen korpuksset ovat vielä verrattain pienikokoisia. Tärkeintä analyysin kannalta kuitenkin on korpusten vertailtavuus, joka on ensisijaista tämänkaltaisessa vertailevassa korpustutkimuksessa. Tästä syystä myös aineiston keruun ja korpusten kokoamisen johdonmukaisuuteen on kiinnitetty erityistä huomiota.

4 Korpusanalyysin tulokset

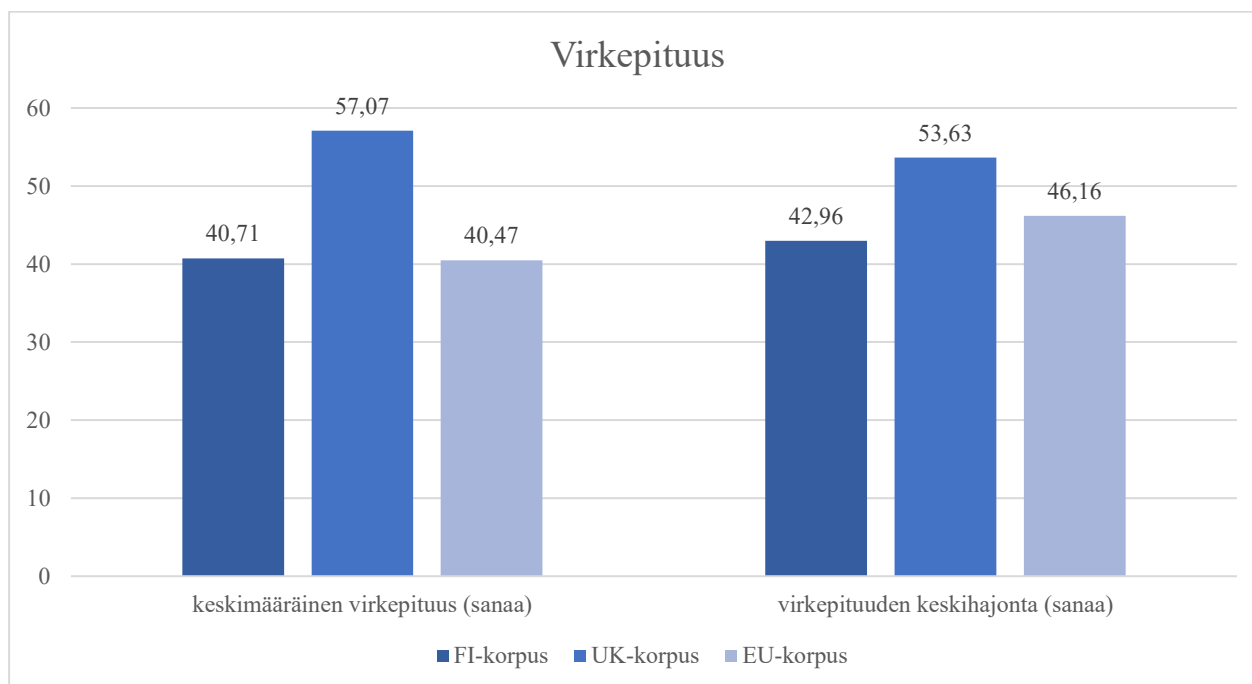
Tässä luvussa esittelen tutkimuksen aineistolle tehdyn korpusanalyysin tulokset. Korpusten analysointiin on käytetty enimmäkseen Sketch Engine -korpustyökalua, mutta virkepituudet ja TTR-luku on laskettu käyttäen apuna Wordsmith-ohjelmaa. Korpusaineistosta on tarkasteltu käännetyn lainsäädännön tekstuaalisten piirteiden yhdenmukaisuutta alun perin kohdekielellä kirjoitettuun lainsäädäntöön. Tekstuaalisia piirteitä havainnoidaan globaalisti makrotasolla virkepituuksien, sanaluokkien, leksikaalisen tiheyden ja TTR-luvun suhteen sekä paikallisesti mikrotasolla sanalistojen, n-grammifrekvenssien sekä juridista logiikkaa kommunikoivien ilmaisujen suhteen. Tavoitteena on näin saavuttaa kokonaiskuva siitä, missä määrin ja millä tavoin suomalaisten lakien englanninkielisistä käännöksistä koostuva FI-korpus eroaa tekstuaalisilta piirteiltään Euroopan unionin asetusten englanninkielisistä toisinoista koostuvasta EU-korpuksesta sekä Yhdistyneen kuningaskunnan parlamentin laeista koostuvasta UK-korpuksesta. Korpusten tekstuaalisen yhdenmukaisuuden tarkastelussa käytän apuna käännösuniversaalihypoteeseja, erityisesti *eksplikoitumista*, *normalisoitumista* sekä *yksinkertaistumista*. Luvun lopuksi esittelen yhteenvedon analyysin tuloksista.

4.1 Tekstuaalinen yhdenmukaisuus makrotasolla

4.1.1 Virkepituudet

Virkepituuksien laskemiseen on käytetty apuna WordSmith-korpustyökalua, sillä Sketch Enginellä ei voi laskea virkepituuksien keskihajontaa. Tulosten tarkastelussa on huomioitava, että virkkeen päättymiseksi käsitetään ainoastaan piste, kysymysmerkki tai huutomerkki, ei rivinvaihtoa tai puolipistettä. Täten eri lakien/asetusten sekä niiden osioiden ja pykäläien/artiklojen numerot ja otsikot luetaan kuuluviksi niitä seuraavaan virkkeeseen, mikä pidentää keskimääräistä virkepituutta. Toisaalta rivinvaihtoa ei voi merkitä virkkeen päättymiseksi, sillä aineiston teksteille on ominaista, että pykälä tai artikla muodostetaan luetelmaksi, jolloin korpustyökalu tulkitsisi virkkeen päättymisen luetelmakohdan vaihtuessa. Muun muassa siksi virkkeet ovat yleiskielisiin teksteihin verrattuna keskimäärin huomattavasti pidempiä, mutta virkkeitä pidentävät esimerkiksi myös oikeuskielelle ominaiset lukuisat kiilalauseet. Tilanne on kaikissa korpuksissa sama, joten otsikoiden ei pitäisi olennaisesti vääristää laskennan tuloksia, kun tarkoituksena on verrata korpuksia keskenään. Jos tämän tutkimuksen korpuksia aiottaisiin verrata muihin korpuksiin, erot annotointikäytännöissä tulisi ottaa silloin huomioon.

Kuvassa 2 on esitetty korpusten keskimääräinen virkepituus pyöristettynä kahden desimaalin tarkkuuteen sekä virkepituuden keskihajonta. FI-korpuksessa on yhteensä 198 955 juoksevaa sanaa 4 887 virkkeessä, ja täten keskimääräiseksi virkepituudeksi saadaan 40,71 sanaa. UK-korpuksessa on 196 368 juoksevaa sanaa 3 441 virkkeessä, ja keskimääräinen virkepituus on 57,07 sanaa. EU-korpuksessa on 196 909 juoksevaa sanaa ja 4 866 virkettä, ja keskimääräinen virkepituus on 40,47 sanaa. Vastaavasti virkepituuden keskihajonta on FI-korpuksessa 42,96 sanaa, UK-korpuksessa 53,63 sanaa ja EU-korpuksessa 46,16 sanaa.



Kuva 2. Keskimääräinen virkepituus ja virkepituuden keskihajonta korpuksista

Yleisesti ottaen korpusten virkkeet ovat oikeuskielelle ominaisesti pitkiä, kiilalauseita sisältäviä ja luettelomaisesti muotoiltuja jättivirkeitä (”periodityylistä” ks. esim. Mattila 2017, 114–115). FI-korpuksen ja EU-korpuksen virkepituudet ovat lähestulkoon identtiset, joskin virkepituuden keskihajonta on EU-korpuksessa 3,2 sanaa korkeampi kuin FI-korpuksessa. P-arvo FI-korpuksen ja EU-korpuksen virkepituuksien keskiarvon erolle on 0,79, mikä tarkoittaa, että on todennäköistä, että ero korpusten välillä johtuu satunnaisesta vaihtelusta. UK-korpuksen keskimääräinen virkepituus sen sijaan on jopa 16,36 sanaa korkeampi kuin FI-korpuksen, ja myös virkepituuden keskihajonta on UK-korpuksessa 10,67 sanaa FI-korpusta korkeampi. FI-korpuksen ja UK-korpuksen välinen keskiarvojen eron P-arvo on pienempi kuin 0,001, joten eron voi katsoa olevan tilastollisesti erittäin merkitsevä. FI-korpuksen virkepituudet ovat siis huomattavasti samankaltaisempia EU-korpuksen kuin UK-korpuksen kanssa.

Virkepituuksien keskihajonta on kaikissa korpuksissa suurta. FI-korpuksessa ja EU-korpuksessa virkepituuksien keskihajonta on virkepitouden keskiarvoa suurempi, mikä tarkoittaa sitä, että molemmissa korpuksissa yli 80 sanan virkkeet ovat yleisiä. Myös UK-korpuksessa keskihajonta on lähestulkoon samansuuruinen kuin keskiarvo, joten korpuksessa on runsaasti myös yli sadan sanan virkeitä. Virkepituuksien keskihajontaa tarkasteltaessa voidaan siis todeta, että vaihtelujen erot korpusten välillä ovat normalisoitumishypoteesin mukaisia. Sekä UK-korpuksessa että EU-korpuksessa virkepituuksien keskihajonta on suurempaa kuin FI-korpuksessa, mikä kielii FI-korpuksen normalisoitumisesta. Toisaalta FI-korpuksen virkepituuksien keskihajonnan ero EU-korpukseen on pieni ja tilastollisesti merkitsemätön, ainoastaan 3,2 sanaa, siinä missä ero UK-korpukseen on 10,67 sanaa.

Korpusten virkepituuDET eivät tue eksplikoitumishypoteesia, mutta eivät myöskään yksinkertaistumishypoteesiin sisältyvää oletusta käännösten lyhyempivirkkeisyydestä, sillä FI-korpus on vain UK-korpusta lyhyempivirkkeinen. Toisaalta institutionaaliset käytännöt tulevat aivan erityisellä tavalla esiin juridisissa teksteissä, ja ne vaikuttavat selkeästi siihen, voiko esimerkiksi pykälä sisältää useamman kuin yhden virkkeen. Lisäksi Bielin mukaan (2014a, 101) käännöksissä pitkiä virkeitä on usein pilkottu osiin, mikä saattaa selittää FI-korpuksen UK-korpukseen nähden lyhyemmät virkkeet, joskin tämä vaatisi lähde-kohde-vertailua, johon tämän tutkimuksen puitteissa ei ole mahdollisuutta. Euroopan unionissa vaaditaan, että lainsäädännön eri kielitoisinnoissa on yhtä monta virkettä (ks. Piehl 2003), mikä saattaa selittää EU-korpuksen virkepituuksien samankaltaisuuden FI-korpukseen nähden.

Todennäköisesti suurin yksittäinen virkepituuksiin vaikuttava tekijä on institutionaaliset laatimiskäytännöt, joiden perusteella pitkälle meneviä johtopäätöksiä eksplikoitumisesta tai yksinkertaistumisesta on mahdotonta tehdä. Toisaalta virkepituuksien analyysi ja erityisesti P-arvojen tarkastelu viittaisivat siihen, että suomalaisten lakien englanninkielisten käännösten laatimisessa säädöksen ulkoasuun liittyvät käytännöt ovat selkeästi lähempänä englanninkielisten EU-asetusten laatimisen käytäntöjä kuin Yhdistyneen kuningaskunnan parlamentin lakien laatimiskäytäntöjä.

4.1.2 Sanaluokat

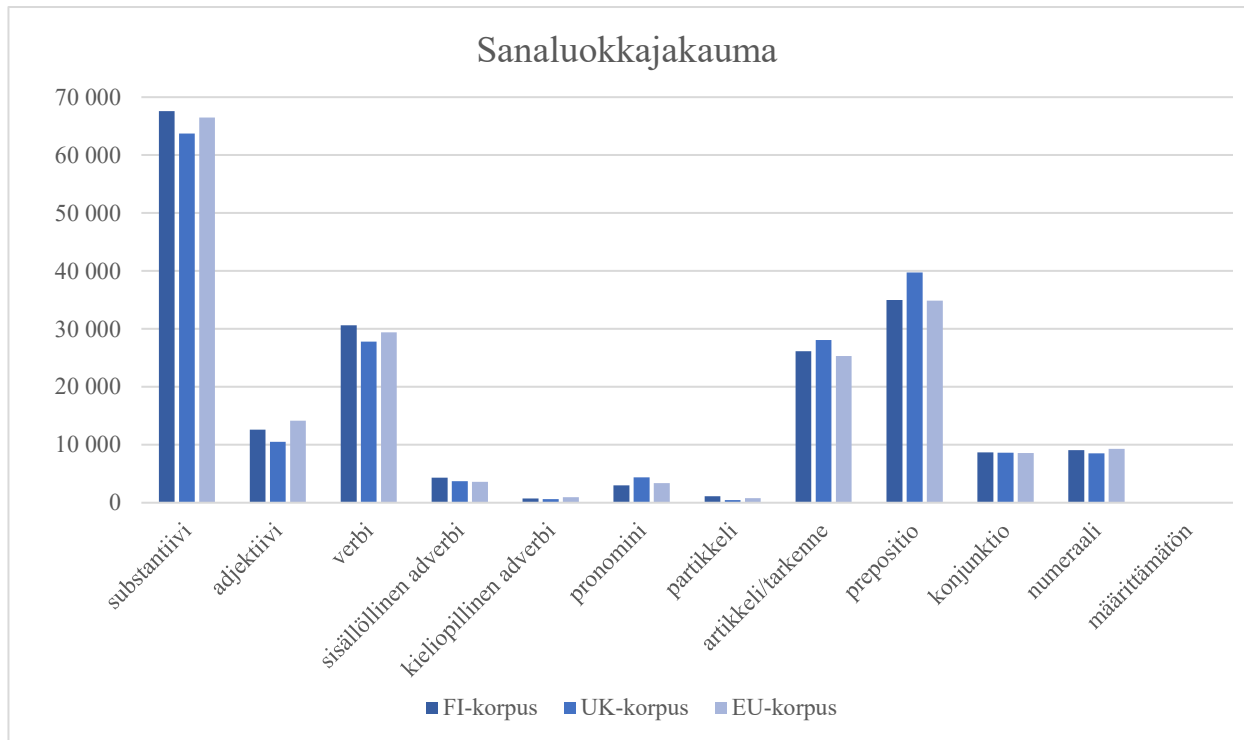
Aineiston sanaluokkajakauma on laskettu sanaluokka-annotoinnin perusteella. Treetaggerin antamat sanaluokat on luokiteltu uudelleen yhteentoista sanaluokkaan (sanaluokat ja frekvenssit liitteessä 2). Treetagger-annotointi on tunnistanut lähestulkoon kaikki korpusten sanat kuuluvaksi johonkin sanaluokkaan (*part of speech, POS*)¹. Tunnistamattomia sanoja on kaikissa korpuksissa yhteenlaskettuna ainoastaan 114 kappaletta (alle 0,02 prosenttia kaikista korpusten sanoista), joten sanaluokkien tunnistus toimii englanniksi huomattavan hyvin samoin kuin sanojen lemmatisaatio. Numeraalien osuus korpuksissa on suhteellisen suuri, sillä esimerkiksi pykälien/artiklojen numeroinnit lasketaan analyysissa numeraaleiksi. Olen kuitenkin poistanut tuloksista sellaiset numerot, jotka Treetagger on havainnut luettelmanumeroinniksi (esim. numerot, joita seuraa rivin alussa piste tai sulkeva sulkumerkki). Numeraalit saattavat hieman vääristää analyysien tuloksia, mutta niitä ei voi kuitenkaan kokonaan poistaa analyysista, sillä niitä ei esiinny ainoastaan otsikoissa ja luotelmissa vaan myös tekstin sisällä, esimerkiksi ristiviittauksissa.

Muuta huomioitavaa sanaluokkien jaottelussa on, että partisiippimuodot on laskettu analyysissa verbeiksi ja että määrittämättömiin sanoihin on sisällytetty tunnistamattomien sanojen lisäksi interjektiot (interjektioita löytyi ainoastaan yksi kappale UK-korpuksesta). FI-korpuksessa määrittämättömiä sanoja on verrattain vähän huolimatta siitä, että teksteissä esiintyy jonkin verran suomenkielisiä erisnimiä. Treetagger on kuitenkin tunnistanut isolla alkukirjaimella kirjoitetut nimet erisnimiksi, ja täten ne on laskettu oikein substantiivien luokkaan. Kuvassa 3 on esitetty diagrammi sanaluokkajakaumista korpuksissa.

Tutkimuksen korpusten sanaluokkajakauma käy sitä eksplikoitumishypoteesiin liittyvää oletusta vastaan, että käännöksissä käytettäisiin selittävien parafraasien takia enemmän konjunktiorakenteita, artikkeleita ja tarkenteita, relatiivipronomineja sekä prepositioita. Konjunktoiden käyttö ei eroa merkittävästi korpusten välillä, ja UK-korpuksessa on pronomineja ja prepositioita huomattavasti enemmän kuin kahdessa muussa korpuksessa. Laajempi pronomien käyttö UK-korpuksessa ja EU-korpuksessa FI-korpukseen verrattuna ja laajempi substantiivien käyttö FI-korpuksessa saattavat sen sijaan kieliä yksinkertaistumiseen liittyvästä toiston lisääntymisestä, joskin erot EU-korpukseen eivät ole tilastollisesti

¹ Sketch Enginen käyttämän Treetagger-annotoinnin annotointitunnisteet ovat nähtävissä internet-sivulla <https://www.sketchengine.eu/english-treetagger-pipeline-2/> [Haettu 18.3.2019].

merkitseviä. Yleisesti ottaen voikin sanoa, että FI-korpuksen ja EU-korpuksen sanaluokkajakaumat korreloivat keskenään, vaikka vaihtelua onkin joissakin sanaluokissa enemmän. Sanaluokkajakaumassa FI-korpuksen erityispiirteenä näkyi huomattavan korkea osuus substantiiveja, verbejä ja sisällöllisiä adverbeja.



Kuva 3. Eri sanaluokkiin kuuluvien sanojen määrät korpuksissa

Eri sanaluokkia edustavien sanojen lukumäärää tarkasteltaessa on huomioitava erot P-arvoissa. FI-korpuksen ja EU-korpuksen substantiivien lukumäärän eron P-arvo on 0,17, kun taas FI-korpuksen ja UK-korpuksen välisen eron P-arvo on pienempi kuin 0,001. FI-korpuksen ja EU-korpuksen välinen substantiivien lukumäärän ero on siis tilastollisesti merkitsemätön siinä missä FI-korpuksen ja UK-korpuksen ero on tilastollisesti erittäin merkitsevä. Toisin sanoen substantiivien käytön taajuus on hyvin samankaltaista FI-korpuksessa ja EU-korpuksessa. Samoin prepositioista lasketut P-arvot eroavat huomattavasti verrokkikorpusten välillä: FI-korpuksen ja UK-korpuksen prepositioiden lukumäärän eron P-arvo on pienempi kuin 0,001, mutta FI-korpuksen ja EU-korpuksen eron P-arvo on peräti 0,35.

Konjunktioiden suhteen korpusten väliset erot eivät ole tilastollisesti merkitseviä: FI-korpuksen ja UK-korpuksen välisen konjunktioiden lukumäärän eron P-arvo on 0,68, ja P-arvo FI-korpuksen ja EU-korpuksen välille laskettuna on 0,77. Myös numeraaleissa vaihtelu on verrattain pientä, joskin molemmista verrokkikorpuksista lasketut P-arvot sijoittuvat 0,001:n ja

0,05:n väliin ja ovat täten joko tilastollisesti melkein merkitseviä tai merkitseviä, mutta eivät erittäin merkitseviä.

Artikkelien ja tarkenteiden suhteen ero FI-korpuksen ja UK-korpuksen välillä P-arvo oli pienempi kuin 0,001 ja täten tilastollisesti erittäin merkitsevä, ja FI-korpuksen ja EU-korpuksen välillä P-arvo oli tilastollisesti merkitsevä 0,004. Muissa sanaluokissa UK-korpus ja EU-korpus eivät eronneet toisistaan tilastollisen merkitsevyyden suhteen, vaan P-arvot näiden korpusten ja FI-korpuksen välillä osoittivat erittäin suurta tilastollista merkitsevyyttä.

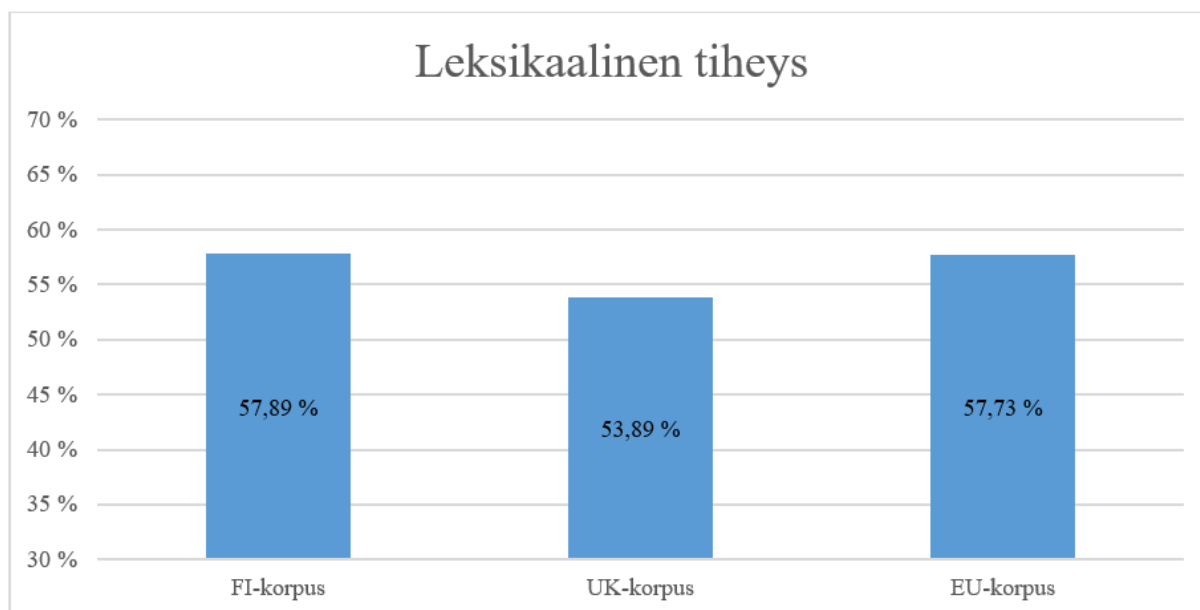
Sanaluokkia tarkasteltaessa FI-korpus näyttäytyy siis huomattavasti samankaltaisempuna EU-korpukseen nähden kuin UK-korpukseen nähden. Sanaluokista pelkästään kieliopillisissa adverbeissa FI-korpus poikkeaa enemmän EU-korpukseen nähden kuin UK-korpukseen nähden (P-arvo FI–UK = 0,10; P-arvo FI–EU < 0,001). Toisaalta vaikka esimerkiksi partikkeleita on FI-korpuksessa 143,1 prosenttia enemmän kuin UK-korpuksessa ja 38,5 prosenttia enemmän kuin EU-korpuksessa, ei adverbien ja partikkelien frekvenssien vaihtelu ole suurta absoluuttisesti, koska partikkeleita ja adverbeja esiintyy korpuksissa melko vähän ylipäänsä.

4.1.3 Leksikaalinen tiheys

Korpusten leksikaalinen tiheys eli sisällöllisten sanojen määrän suhde kaikkiin korpuksen sanoihin on laskettu jakamalla sisällöllisten sanojen lukumäärä koko korpuksen sanamäärällä. Sisällöllisinä sanoina on otettu huomioon substantiivit, adjektiivit, verbit sekä sisällölliset adverbit. Sisällöllisiksi adverbeiksi on sisällytetty Treetaggerin annotointitunnisteet RB (adverbien yleisluokka, esim. *negligently, jointly, above, otherwise*), RBR (tapaa ilmaiseva komparatiivimuotoinen adverbi, esim. *quicker, better*) sekä RBS (tapaa ilmaiseva superlatiivimuotoinen adverbi, esim. *best, fastest*). Kieliopillisiksi adverbeiksi luettiin luokat WRB (wh-adverbit kuten *when* ja *where*) ja EX (sana *there* eksistentiaalilauseessa). Osa luokkaan RB kuuluvista adverbeista on kieliopillisia (esim. *not, only*), mutta niiden erottaminen aineistosta olisi työlästä, sillä Treetagger ei erottele vaadittavalla tarkkuudella sisällöllisiä ja kieliopillisia adverbeja, ja myös manuaalinen erottelu vaatisi selvän rajanvedon ja olisi työlästä. Korpusten leksikaaliset tiheydet on kokeiltu laskea myös lukien sisältösanoiksi ainoastaan substantiivit, adjektiivit ja verbit, ja tällöinkään korpuksien keskinäinen leksikaalisen tiheyden suhde ei huomattavasti muutu. Lisäksi adverbien määrä ylipäänsä korpuksissa on suhteellisen pieni – FI-korpuksessa 2,5 % saneista, UK-korpuksessa 2,2 % ja EU-korpuksessa 2,3 % – joten voi olettaa, että adverbien erottelu on tämän tutkimuksen tarkoituksiin ollut riittävän tarkka

ottaen huomioon, että jakoperuste on kaikkien korpusten kohdalla ollut sama ja että korpuksia verrataan tutkimuksessa ainoastaan toisiinsa. Korpusten leksikaaliset tiheydet on esitetty alla kuvassa 4, ja eri sanaluokkiin sisällytettyjen sanojen lukumäärät ovat nähtävillä liitteessä 2.

Leksikaalinen tiheys LT on siis laskettu kaavalla $LT = \frac{N_{lex}}{N}$, jossa N_{lex} tarkoittaa sisällöllisten sanojen määrää korpuksessa ja N kaikkien sanojen yhteenlaskettua määrää korpuksessa. Näin saadaan tulokseksi leksikaalisen tiheyden suhdeluku.



Kuva 4. Korpusten leksikaaliset tiheydet

FI-korpuksessa sisältösanoja on yhteensä 115 180 kappaletta, ja juoksevia sanoja korpuksessa on yhteensä 198 955, joten leksikaalinen tiheys $LT = \frac{115\,180}{198\,955} = 0,5789 \rightarrow 57,89\%$.

UK-korpuksessa sisältösanoja on 105 826, joten juoksevien sanojen määrän ollessa 196 368 leksikaalinen tiheys $LT = \frac{105\,826}{196\,368} = 0,5389 \rightarrow 53,89\%$.

EU-korpuksessa sisältösanoja on yhteensä 113 673 kappaletta, ja juoksevia sanoja korpuksessa on yhteensä 196 909, joten leksikaalinen tiheys $LT = \frac{113\,673}{196\,909} = 0,5773 \rightarrow 57,73\%$.

FI-korpuksen leksikaalinen tiheys on kolmesta korpuksesta suurin. Toisaalta FI-korpuksen leksikaalinen tiheys on ainoastaan 0,16 prosenttiyksikköä suurempi kuin EU-korpuksen. Ero EU-korpuksen ei ole tilastollisesti merkitsevä: P-arvoksi laskettiin 0,30. UK-korpuksen leksikaalinen tiheys sen sijaan on huomattavasti muita korpuksia pienempi ja FI-korpuksen

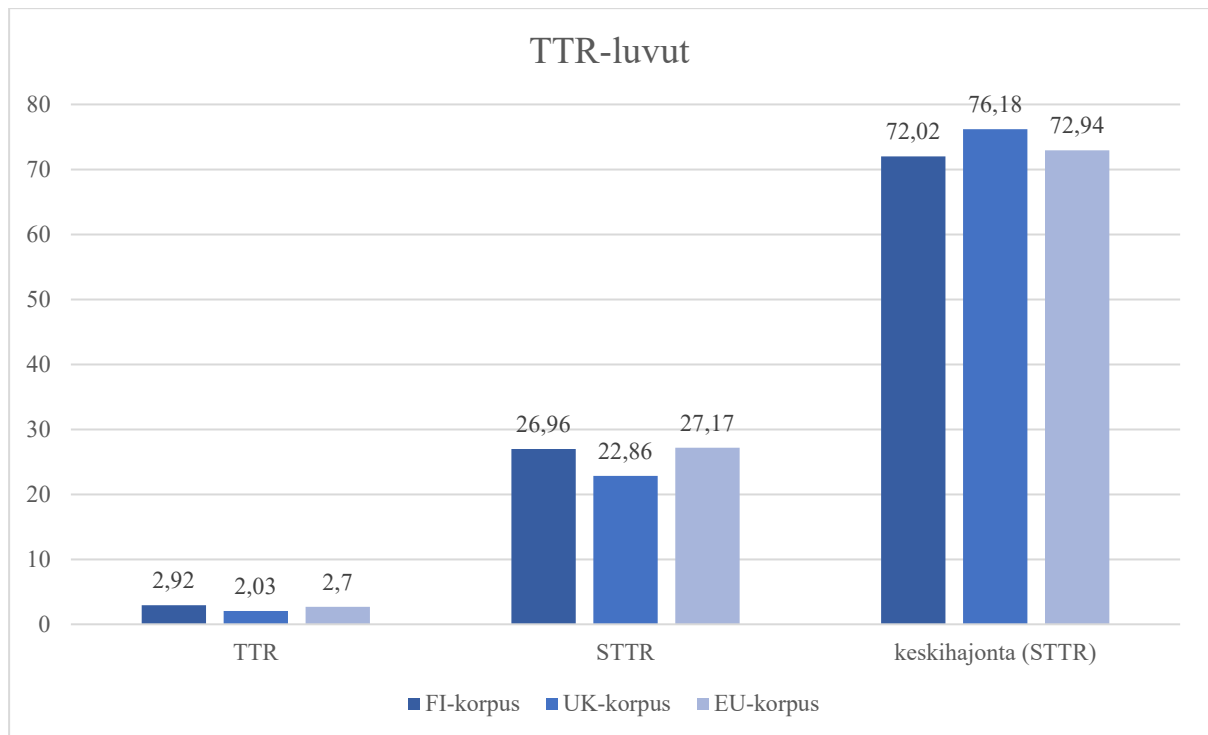
eroa on 4 prosenttiyksikköä, ja FI-korpuksen ja UK-korpuksen erolle laskettu P-arvo on pienempi kuin 0,001.

FI-korpus on siis leksikaalisen tiheyden suhteen huomattavasti lähempänä EU-korpusta kuin UK-korpusta. FI-korpuksen ja EU-korpuksen välinen suhteellisen suuri ja tilastollista merkitsemättömyyttä ilmaiseva P-arvo kielii siitä, että nämä kaksi korpusta ovat keskenään samankaltaisia leksikaalisen tiheyden suhteen, kun taas olisi erittäin epätodennäköistä, että FI-korpuksen ja UK-korpuksen välinen ero leksikaalisessa tiheydessä johtuisi satunnaisesta vaihtelusta. Yksinkertaistumishypoteesin mukaan käännösten leksikaalisen tiheyden tulisi olla pienempi kuin alun perin kohdekielellä kirjoitettujen tekstien, mutta analyysin tulokset ovat ristiriidassa tämän hypoteesin kanssa.

4.1.4 TTR-luku

TTR-luku eli *sana-sane-suhde* tarkoittaa yksittäisten sanojen suhdetta kaikkien saneiden lukumäärään (engl. *type-token ratio*). Korpuksen TTR-luku lasketaan jakamalla yksittäisten sananmuotojen lukumäärä tekstissä koko korpuksen saneiden lukumäärällä. TTR-arvon laskemiseen on käytetty Wordsmith-korpustyökalua. TTR on hyvä apuväline toiston mittaamiseen yksittäisissä teksteissä, mutta suurista tekstimassoista kuten korpuksista saatava TTR-luku on väistämättä hyvin pieni ja riippuvainen tekstien pituudesta, minkä vuoksi Wordsmith laskee tavallisen TTR-luvun lisäksi myös standardisoidun TTR-arvon tuhannen sanan välein. Korpusten TTR-luvut on esitetty alla kuvassa 5.

FI-korpuksessa TTR-luku on 2,92 prosenttia, UK-korpuksessa 2,03 prosenttia ja EU-korpuksessa 2,70 prosenttia. Toisin sanoen esimerkiksi FI-korpuksessa on sataa sanetta kohti 2,92 yksittäistä sanaa. UK-korpuksen pienempään TTR-lukuun todennäköisesti vaikuttaa se, että korpuksessa on vähemmän yksittäisiä tekstejä kuin muissa korpuksissa. Toisin sanoen jotkut UK-korpuksen tekstit ovat huomattavasti pidempiä kuin toiset, jolloin myös toistoa voi olettaa esiintyvän enemmän. Standardisoimattoman TTR-luvun onkin todettu olevan pitkälti riippuvainen tekstin pituudesta (ks. esim. Biel 2014a, 139), minkä vuoksi standardisoitu TTR-luku on parempi mittari korpusten vertailuun.



Kuva 5. Korpusten TTR-luvut

Tuhannen sanan jaksoissa laskettu standardisoitu TTR-luku (engl. *standardized type-token ratio*, *STTR*) on FI-korpuksessa 26,96 prosenttia, UK-korpuksessa 22,86 prosenttia ja EU-korpuksessa 27,17 prosenttia. STTR-luku on sen vuoksi standardisoimatonta TTR-lukua huomattavasti korkeampi, että tuhannen sanan jaksoissa samat sananmuodot eivät ehdi toistumaan yhtä useasti kuin koko korpusta tarkasteltaessa. Standardisoidun TTR-luvun keskihajonta on FI-korpuksessa 72,02 prosenttiyksikköä, UK-korpuksessa 76,18 prosenttiyksikköä ja EU-korpuksessa 72,94 prosenttiyksikköä. FI-korpuksen ja UK-korpuksen STTR-luvun eron P-arvoksi laskettiin 0,58 ja FI-korpuksen ja EU-korpuksen väliseksi P-arvoksi 0,98, joten kummankin verrokkikorpuksen kohdalla on todennäköistä, että ero johtuu satunnaisesta vaihtelusta. Eri korpusten välillä ei siis ole huomattavaa eroa standardisoituja TTR-lukuja tarkasteltaessa, mutta FI-korpus on P-arvoja tarkasteltaessa kuitenkin jokseenkin samankaltaisempi EU-korpuksen kuin UK-korpuksen kanssa.

UK-korpuksessa TTR-luku on jonkin verran pienempi, mutta ero saattaa johtua osittain siitä, että UK-korpuksessa on lukumääräisesti vähemmän tekstejä kuin kahdessa muussa korpuksessa, joten pitkälle meneviä johtopäätöksiä TTR-luvuista ei voida vetää. Toisaalta standardisoituun TTR-lukuun sillä ei pitäisi olla vaikutusta, onko korpuksessa pitkiä vai lyhyitä tekstejä. UK-korpuksessa on toistoa silti enemmän myös tuhannen saneen osiin jaettuna, joten tulokset eivät tue yksinkertaistumishypoteesia, jonka mukaan käänöksissä olisi enemmän

toistoa kuin alun perin kohdekielellä kirjoitetuissa teksteissä. Keskihajonta on hieman korkeampi UK-korpuksessa kuin muissa korpuksissa, joten vaihtelu standardisoidun TTR-luvun laskemiseen käytettyjen tuhannen sanan jaksojen välillä on suurempaa kuin muissa korpuksissa. Standardisoidun TTR-luvun keskihajonnat ovatkin jokseenkin linjassa normalisoitumishypoteesin kanssa. Sekä UK-korpuksessa että EU-korpuksessa STTR-arvon keskihajonta on suurempaa kuin FI-korpuksessa, mikä kielii FI-korpuksen normalisoitumisesta. Toisaalta erot ovat pieniä ja lasketut P-arvot korkeita, joten erot saattavat johtua satunnaisesta vaihtelusta.

4.2 Tekstuaalinen yhdenmukaisuus mikrotasolla

4.2.1 Sanalistat

Korpusten sanastoa on analysoitu tutkimuksessa sanalistojen ja konkordanssihakujen avulla. Sanastosta on pyritty löytämään yhtäläisyyksiä ja eroja korpusten välillä taajaan esiintyvien sanojen käytössä. Suurin osa taajimpaan esiintyvistä sanoista on kaikissa kolmessa korpuksessa funktiosanoja kuten artikkeleita, prepositioita ja konjunktioita.

Sanalistoihin ajetut sanat eivät ole lemmoja vaan yksittäisiä sananmuotoja, joten sanojen eri taivutusmuodot, esimerkiksi eri aikamuodot ja sijamuodot, on laskettu analyysissa omiksi sanoikseen. Näin voidaan havainnoida paremmin sanojen käyttöä sellaisena kuin ne teksteissä esiintyvät. 200 kertaa tai sitä useammin esiintyviä sanoja on FI-korpuksessa 121 kappaletta, UK-korpuksessa 129 kappaletta ja EU-korpuksessa 125 kappaletta, ja 1000 kertaa tai sitä useammin esiintyviä sanoja on FI-korpuksessa 25 kappaletta, UK-korpuksessa 29 kappaletta ja EU-korpuksessa 18 kappaletta.

Olen verrannut korpuksien välillä suuren frekvenssin sanojen (1000 esiintymää tai enemmän) esiintymien lukumääriä, ja taulukossa 2 on esitetty FI-korpuksesta ajetun sanalistan suurifrekvenssisten sanojen frekvenssit muissa korpuksissa (UK-korpuksesta ja EU-korpuksesta ajetut sanalistat liitteessä 3). Kuvassa 6 on esitetty tuhat kertaa ja sitä enemmän esiintyvien sanojen yhteenlasketut esiintymät korpuksissa. Koska tämän tutkimuksen korpuksat ovat lähestulkoon yhtä suuria keskenään, on esiintymien lukumäärät esitetty absoluuttisina sen sijaan, että ne olisi suhteutettu esimerkiksi miljoonaan sanaan. UK-korpus on sanamääräisesti vain 0,12 % suurempi kuin FI-korpus ja EU-korpus vastaavasti 0,11 % sitä pienempi, ja lisäksi korpusten eroille lasketuissa P-arvoissa huomioidaan korpusten kokoerot, joten lukumäärien suhteuttaminen ei ole tarpeen.

Yleisesti ottaen voi todeta, että sanalistoissa esiintyvät funktiosanat ovat tyypillisesti artikkeleita, prepositioita ja konjunktioita. Viisi useimmin esiintyvää FI-korpuksen ja EU-korpuksen sanaa on samassa järjestyksessä, ja myös UK-korpuksen kolme eniten esiintyvää sanaa on samoja. Artikkelin *the* jälkeen kaikissa korpuksissa useimmiten esiintyy *of*-prepositio, mikä kielii oikeuskielelle ominaiseen tapaan *of*-genetiivin käytöstä possessiivisuffiksin *s* sijaan. *S*-possessiivisuffiksi löytyykin konkordanssihaulla CQL-syntaksia [*tag*= "NNZ|NNSZ"] käyttäen FI-korpuksesta ainoastaan 225 kertaa, UK-korpuksesta 199 kertaa ja EU-korpuksessa 180 kertaa, kun taas *of* esiintyy FI-korpuksessa 9 064 kertaa, UK-korpuksessa 10 829 kertaa ja EU-korpuksessa 9 603 kertaa.

Taulukko 2. FI-korpuksessa vähintään 1000 kertaa esiintyvien sanojen esiintymät muissa korpuksissa

sijoitus FI-korpuksessa	sana	esiintymät korpuksissa			P-arvo	
		FI-korpus	UK-korpus	EU-korpus	FI-korpus–UK-korpus	FI-korpus–EU-korpus
1	the	19 006	14 089	17 211	P < 0,001	P < 0,001
2	of	9 064	10 829	9 603	P < 0,001	P < 0,001
3	to	5 759	7 282	6 292	P < 0,001	P < 0,001
4	in	5 395	5 635	5 270	P = 0,003	P = 0,49
5	and	5 273	3 394	5 248	P < 0,001	P = 0,77
6	a	3 857	6 405	3 347	P < 0,001	P < 0,001
7	or	3 360	5 052	3 210	P < 0,001	P = 0,15
8	for	2 996	2 623	2 686	P < 0,001	P < 0,001
9	section	2 582	2 522	35	P = 0,71	P < 0,001
10	be	2 451	2 265	2 079	P = 0,02	P < 0,001
11	on	2 449	1 100	1 415	P < 0,001	P < 0,001
12	act	1 940	1 587	74	P < 0,001	P < 0,001
13	is	1 868	2 029	870	P = 0,003	P < 0,001
14	shall	1 778	1 050	3 872	P < 0,001	P < 0,001
15	by	1 561	2 826	1 642	P < 0,001	P = 0,08
16	as	1 255	1 942	1 239	P < 0,001	P = 0,95
17	this	1 202	1 745	1 098	P < 0,001	P = 0,051
18	referred	1 164	73	753	P < 0,001	P < 0,001
19	that	1 143	1 852	1 467	P < 0,001	P < 0,001
20	waste	1 137	243	405	P < 0,001	P < 0,001
21	may	1 132	1 583	710	P < 0,001	P < 0,001
22	provisions	1 131	286	252	P < 0,001	P < 0,001
23	if	1 040	568	386	P < 0,001	P < 0,001
24	with	1 040	1 131	1 796	P = 0,02	P < 0,001
25	an	1 031	1 348	859	P < 0,001	P < 0,001

Which-relatiivipronominin odotettiin esiintyvän taajaan FI-korpuksessa, mutta *which* esiintyy UK-korpuksessa huomattavasti taajempaan kuin muissa korpuksissa, 1 507 kertaa, kun taas FI-

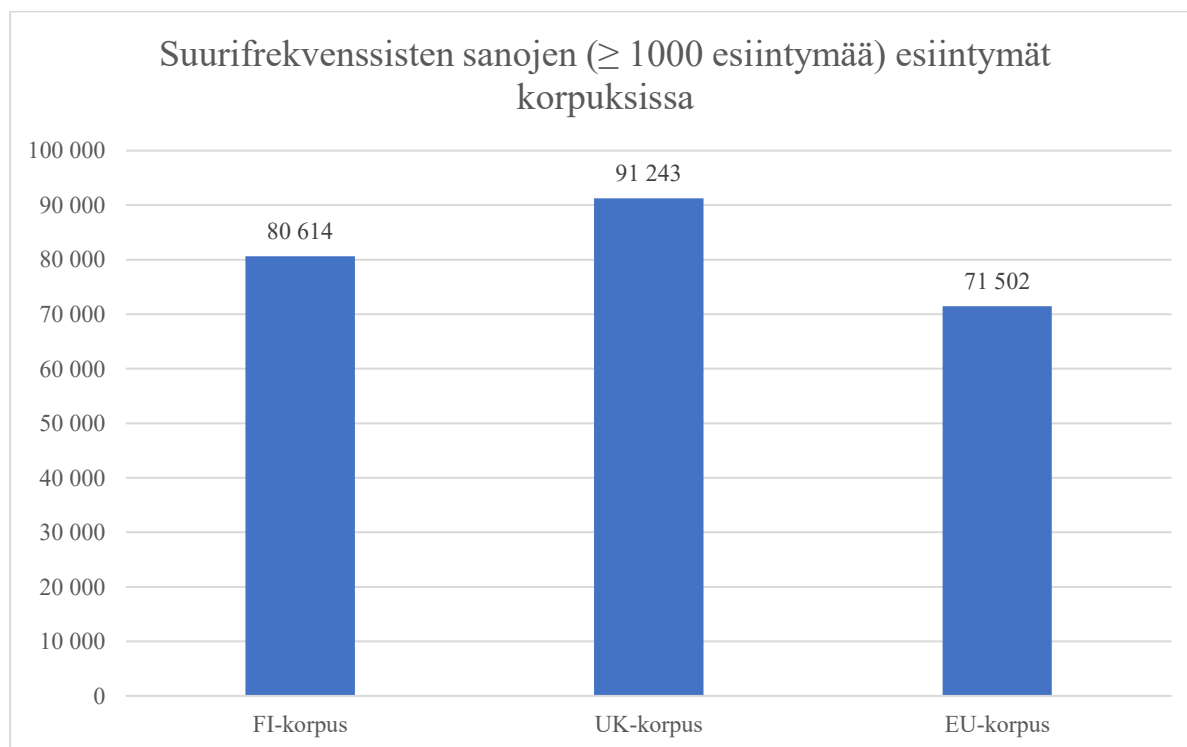
korpuksessa ja EU-korpuksessa pelkästään 543 ja 800 kertaa. Eksplikoitumista lisäävissä selittävissä parafraaseissa käytetään usein *which*-pronominia, mutta kyseistä pronominia ei näy kuitenkaan käytettävän FI-korpuksessa yhtä paljon kuin alun perin kohdekielellä kirjoitetuissa teksteissä. Pronominin *which* esiintyminen käännöskorpuksissa ei siis tue eksplikoitumishypoteesia selittävien parafraasien osalta.

Huomionarvoista on myös se, että rinnastuskonjunktia *or* on käytetty enemmän UK-korpuksessa kuin muissa korpuksissa mutta taas konjunktia *and* vähemmän. *Or* esiintyy FI-korpuksessa 3 360 kertaa, UK-korpuksessa 5 052 kertaa ja EU-korpuksessa 3 210 kertaa; *and* esiintyy FI-korpuksessa 5 273 kertaa, UK-korpuksessa 3 394 kertaa ja EU-korpuksessa 5 248 kertaa. Täten *or* esiintyy FI-korpuksessa 33,49 prosenttia vähemmän kuin UK-korpuksessa ja 4,67 prosenttia enemmän kuin EU-korpuksessa. FI-korpuksen ja UK-korpuksen ero *or*-sanankäytössä on tilastollisesti erittäin merkitsevä ($P < 0,001$) ja FI-korpuksen ja EU-korpuksen ero taas ei ($P = 0,15$). *And* puolestaan esiintyy FI-korpuksessa 55,36 prosenttia enemmän kuin UK-korpuksessa ja 0,48 prosenttia enemmän kuin EU-korpuksessa, ja FI-korpuksen ja UK-korpuksen ero on jälleen tilastollisesti erittäin merkitsevä ($P < 0,001$), kun taas FI-korpuksen ja EU-korpuksen ero saattaa johtua satunnaisesta vaihtelusta ($P = 0,77$). Rinnastuskonjunktioiden suhteen FI-korpus on siis huomattavasti lähempänä EU-korpusta kuin UK-korpusta. On kuitenkin huomattava, että konjunktilla *or* on erittäin tarkka juridinen funktio lainsäädännöllisissä teksteissä, kun taas konjunktia *and* saatetaan käyttää monenlaisissa yhteyksissä, mikä saattaa selittää sen suurempaa käyttöä FI-korpuksessa. On kuitenkin mielenkiintoista, että konjunktia *and* ei käytetä laajemmin UK-korpuksessa johtuen common law -englantiin juurtuneista sanapareista, kuten *devise and bequeath* tai *goods and chattels* (ks. Mattila 2017, 417).

Määräinen artikkeli *the* esiintyy FI-korpuksessa 19 006 kertaa, UK-korpuksessa 14 089 kertaa ja EU-korpuksessa 17 211 kertaa. Epämääräiset artikkelit *a/an* sen sijaan esiintyvät yhteensä FI-korpuksessa 4 888 kertaa, UK-korpuksessa 7 753 kertaa ja EU-korpuksessa 4 206 kertaa. FI-korpuksessa siis artikkeli *the* esiintyy 34,90 prosenttia enemmän kuin UK-korpuksessa ja 10,43 prosenttia enemmän kuin EU-korpuksessa; artikkelit *a/an* esiintyvät laskettuna yhteen FI-korpuksessa 36,95 prosenttia vähemmän kuin UK-korpuksessa ja 16,21 prosenttia enemmän kuin EU-korpuksessa. FI-korpus ei ole artikkeleiden käytön suhteen selkeästi lähempänä kumpaakaan verrokkikorpusta; erot ovat tilastollisesti erittäin merkitseviä molempiin suuntiin ($P < 0,001$). Tarkenteen *this* käyttö FI-korpuksessa sen sijaan on lähempänä EU-korpusta kuin

UK-korpusta. Laskettu P-arvo erosta UK-korpukseen on pienempi kuin 0,001, kun taas erolle EU-korpukseen laskettiin P-arvo 0,051, joka on juuri ja juuri tilastollisen merkitsevyyden ulkopuolella.

Myös prepositioiden käytössä FI-korpus seuraa jokseenkin EU-korpusta. Prepositio *in* esiintyy FI-korpuksessa 5 395 kertaa, UK-korpuksessa 5 635 kertaa ja EU-korpuksessa 5 270 kertaa. FI-korpuksen ja UK-korpuksen välinen ero on tilastollisesti merkitsevä ($P = 0,003$) ja FI-korpuksen ja EU-korpuksen välinen ei ($P = 0,49$). Samoin preposition *to* käyttö FI-korpuksessa on lähempänä EU-korpusta, vaikkakin ero myös EU-korpukseen on tilastollisesti merkitsevä, ja preposition *by* ero UK-korpukseen on tilastollisesti erittäin merkitsevä ($P < 0,001$) ja ero EU-korpukseen ei ($P = 0,08$). Sanan *as* käyttö FI-korpuksessa on myös selvästi lähempänä EU-korpusta eron P-arvon ollessa peräti 0,95, kun taas UK-korpukseen eroa on selvästi enemmän ja P-arvo on pienempi kuin 0,001. Sen sijaan preposition *with* käyttö seuraa pikemmin UK-korpusta: ero UK-korpukseen on tilastollisesti melkein merkitsevä ($P = 0,02$), mutta P-arvo on huomattavasti suurempi kuin verrattaessa FI-korpusta EU-korpukseen ($P < 0,001$).



Kuva 6. Tuhat kertaa tai enemmän esiintyvien sanojen esiintymien lukumäärät korpuksissa

UK-korpuksen kanssa yhdenmukaisempia sanoja ovat myös substantiivi *section*, joka esiintyy FI-korpuksessa 2 582 kertaa, UK-korpuksessa 2 522 kertaa ja EU-korpuksessa pelkästään 35

kertaa, sekä verbi *be*, joka esiintyy FI-korpuksessa 2 451 kertaa, UK-korpuksessa 2 265 kertaa ja EU-korpuksessa 2 079 kertaa. *Section*-sanon ero FI-korpuksen ja UK-korpuksen välillä on todennäköisesti seurausta satunnaisesta vaihtelusta ($P = 0,71$), kun taas ero EU-korpukseen on tilastollisesti erittäin merkitsevä ($P < 0,001$). *Be*-sanon ero FI-korpuksen ja UK-korpuksen välillä on tilastollisesti melkein merkitsevä ($P = 0,02$), kun taas ero EU-korpukseen on erittäin merkitsevä ($P < 0,001$).

FI-korpus on siis sanalistoja tarkasteltaessa jokseenkin tekstuaalisesti yhdenmukaisempi EU-korpuksen kuin UK-korpuksen kanssa, mutta vaihtelu on suurta, ja suurimmassa osassa vertailuista lukumääristä eron laskettu P -arvo on pienempi kuin tilastollisesti erittäin merkitsevän eron raja. Esimerkiksi performatiivia *shall* käytetään EU-korpuksen teksteissä erityisen paljon (*shallista* lisää alaluvussa 4.2.3). Suuri osuus taajaan esiintyviä sanoja tarkoittaisi Chestermanin mukaan yksinkertaistumista (Chesterman 2010, 177), mutta yli tuhannen esiintymän yksiköitä on huomattavasti eniten UK-korpuksessa, ja UK-korpuksen tuhat kertaa tai enemmän esiintyvät sanat muodostavat myös huomattavasti suuremman osan korpuksen kaikista juoksevista sanoista kuin kahdessa muussa korpuksessa: UK-korpuksessa nämä sanat muodostavat 46,47 prosenttia kaikista korpuksen juoksevista sanoista, FI-korpuksessa 40,52 prosenttia ja EU-korpuksessa 36,31 prosenttia. Erot ovat myös tilastollisesti erittäin merkitseviä ($P < 0,001$). Tulokset ovat siis ristiriidassa yksinkertaistumishypoteesin kanssa. UK-korpuksen pienemmän lukumäärän yksittäisiä tekstejä pitäisi vaikuttaa ainoastaan sisällöllisten sanojen esiintymiin, ei niinkään funktiosanoihin.

4.2.2 N-grammifrekvenssit

N-grammien frekvenssien analysoinnissa on otettu huomioon kaikki 3–6 sanan mittaiset sanajonot, jotka esiintyvät korpuksessa vähintään 50 kertaa. N-grammifrekvensseistä verrataan eroja korpuksissa usein toistuvien fraasien rakenteissa ja sisällössä. 50 kertaa tai sitä useammin esiintyviä n-grammeja oli FI-korpuksessa 218 kappaletta, UK-korpuksessa 248 kappaletta ja EU-korpuksessa 177 kappaletta. Taulukossa 3 on esitetty FI-korpuksessa suuren frekvenssin n-grammien (200 esiintymää tai enemmän) esiintymät muissa korpuksissa (UK-korpuksesta ja EU-korpuksesta ajatut n-grammilistat liitteessä 4).

Hieman yllättäen korpusten useimmiten esiintyvien n-grammien joukosta ei löytynyt niinkään oikeudellisia fraseologismeja, vaan taajaan esiintyvät n-grammit liittyvät pikemmin ristiviittauksiin ja tiedon jäsentelyyn säädöksissä. Suurifrekvenssisten n-grammien

lukumäärässä ei ole suurta eroa korpusten välillä, mutta n-grammeissa itsessään on huomattavia eroja. N-grammien erot ovat lähes yksinomaan tilastollisesti erittäin merkitseviä, minkä vuoksi korpusten vertailu P-arvojen perusteella on hankalaa. Pelkästään ilmauksissa *referred to in* sekä *in accordance with* havaittiin eroa FI-korpuksen ja UK-korpuksen sekä FI-korpuksen ja EU-korpuksen välisessä tilastollisessa merkitsevyydessä. Yleisesti ottaen voidaan löytää yhtäläisyyksiä FI-korpuksen ja EU-korpuksen välillä, sillä monia FI-korpuksessa suurifrekvenssisiä n-grammeja esiintyy runsaasti myös EU-korpuksessa ja toisin päin, mutta UK-korpuksessa muiden korpusten suurifrekvenssisiä n-grammeja ei juuri esiinny eikä UK-korpuksen suurifrekvenssisiä n-grammeja muissa korpuksissa.

Taulukko 3. FI-korpuksen suuren frekvenssin n-grammien (≥ 200 esiintymää) esiintymät muissa korpuksissa

sijoitus FI-korpuksessa	n-grammi	esiintymät korpuksissa			P-arvo	
		FI-korpus	UK-korpus	EU-korpus	FI-korpus–UK-korpus	FI-korpus–EU-korpus
1	referred to in	1 112	37	730	P < 0,001	P = 0,06
2	to in section	518	4	1	P < 0,001	P < 0,001
3	referred to in section	516	4	1	P < 0,001	P < 0,001
4	laid down in	449	3	156	P < 0,001	P < 0,001
5	of this act	290	129	0	P < 0,001	P < 0,001
6	in accordance with	266	236	830	P = 0,05	P < 0,001
7	the act on	265	0	0	P < 0,001	P < 0,001
8	provisions on the	251	0	3	P < 0,001	P < 0,001
9	the ministry of	240	0	0	P < 0,001	P < 0,001
10	and the environment	222	2	6	P < 0,001	P < 0,001
11	to in subsection	201	22	0	P < 0,001	P < 0,001
12	referred to in subsection	200	22	0	P < 0,001	P < 0,001
13	in this act	200	28	0	P < 0,001	P < 0,001

EU-korpuksessa ilmaus *pursuant to article* esiintyy 394 kertaa, kun taas FI-korpuksessa ja UK-korpuksessa ainoastaan 2 ja 3 kertaa, eikä sanaa *pursuant* näy käytettävän myöskään ilman sanaa *article*: konkordanssihaulla sana *pursuant* esiintyy FI-korpuksessa yhteensä 62 kertaa, UK-korpuksessa 51 kertaa ja EU-korpuksessa 680 kertaa. *Pursuant to* -rakennetta käytetään siis huomattavan paljon enemmän EU-korpuksen teksteissä kuin kahdessa muussa korpuksessa.

Erityisesti rakenne *referred to in* esiintyy huomattavan paljon FI-korpuksessa, 1 112 kertaa, kun taas EU-korpuksessa se esiintyy 730 kertaa ja UK-korpuksessa pelkästään 37 kertaa. Tässä mielessä FI-korpus on selvästi yksinkertaistunut UK-korpuksen nähden, jossa samaa funktiota

ilmaiseva rakenne muodostetaan useammilla tavoilla (esim. *mentioned in, under, specified in*). Toisaalta rakenne *in accordance with* esiintyy lähestulkoon samassa suhteessa FI-korpuksessa ja UK-korpuksessa (266 ja 236 esiintymää, eron P-arvo = 0,05), mutta EU-korpuksessa kyseinen n-grammi on selvästi ylliedustunut (830 esiintymää, eron P-arvo < 0,001), joten yksinkertaistumista todennäköisempää lienee, että erot johtuvat käännösprosessin ulkoisista tekijöistä, kuten institutionaalisista tekstin laatimisen käytännöistä. Yksinkertaistumishypoteesin mukaan käännöksissä myös suurifrekvenssisia yksiköitä tulisi ylipäänsä olla enemmän: vähintään 200 esiintymän n-grammeja on FI-korpuksessa 13 kappaletta, UK-korpuksessa 12 kappaletta ja EU-korpuksessa 18 kappaletta, ja vähintään 50 esiintymän n-grammeja on FI-korpuksessa 218 kappaletta, UK-korpuksessa 248 kappaletta ja EU-korpuksessa 177 kappaletta. Havaitut määrät eivät siis tue yksinkertaistumishypoteesia.

4.2.3 Juridisen logiikan kommunikointi

Juridisen logiikan kommunikointiin liittyvän sanaston analyysissä tarkasteltiin oikeuskielessä käytettäviä ehtolauseilmauksia sekä oikeutta, velvollisuutta, lupaa ja kieltoa kommunikoivia performatiiveja. Analyysi suoritettiin vertaamalla konkordanssihaulla tarkastelluista yksiköistä saatujen frekvenssien eroja korpusten välillä.

Tarkastellut ehtolauseilmaukset on valittu Bielin (2014a, 144) esittämän listauksen pohjalta, johon on lisätty ilmaukset *if, who* ja *on (the) condition*. Ehtolauseilmauksina on siis tarkasteltu seuraavia ilmauksia: *apart from, except, in so far as / insofar as, in the absence of, in (the) case(s), in the event, notwithstanding, provided that, save for, unless, when, whenever, where, if, who* sekä *on (the) condition*. Ilmauksien esiintymiä on verrattuna konkordanssihaulla eri korpuksissa, ja tulokset on esitetty taulukossa 4. Tulosten tarkastelussa on huomioitava, että korpuksissa saattaa esiintyä myös muita tähän analyysiin sisällyttämättömiä ilmauksia, joilla juridisia ehtolauseita ilmaistaan. Sanat *when* ja *where* saattavat lisäksi esiintyä myös muissa yhteyksissä kuin ehtolauseissa, mutta silmämääräisellä tarkastelulla esiintymät liittyvät selvästi enimmäkseen ehtolauseisiin.

Ehtolauseilmauksissa on runsaasti vaihtelua eri korpusten välillä. FI-korpuksessa kuitenkin konjunktio *if* on huomattavan ylliedustunut kahteen muuhun korpukseen verrattuna: FI-korpuksessa *if* esiintyy 1 040 kertaa, UK-korpuksessa 568 kertaa ja EU-korpuksessa 386 kertaa. *If* esiintyy FI-korpuksessa siis 83,1 prosenttia enemmän kuin UK-korpuksessa ja 169,43 prosenttia enemmän kuin EU-korpuksessa. Ero molempiin korpuksiin on tilastollisesti erittäin merkitsevä (P < 0,001). *When*-konjunktio esiintyy FI-korpuksessa 386 kertaa, UK-korpuksessa

94 kertaa ja EU-korpuksessa 204 kertaa. *When* esiintyy siis FI-korpuksessa 310,64 prosenttia enemmän kuin UK-korpuksessa ja 89,22 prosenttia enemmän kuin EU-korpuksessa. Konjunktion *when* esiintyminen FI-korpuksessa on siis huomattavasti lähempänä EU-korpusta kuin UK-korpusta, joskin ero molempiin verrokkikorpuksiin on tilastollisesti erittäin merkitsevä ($P < 0,001$). Mielenkiintoista on, että EU-korpuksessa *where* on jopa yleisempi kuin *if*, ja *where* esiintyy EU-korpuksessa myös huomattavasti enemmän kuin kummassakaan muussa korpuksessa. *Where* esiintyy FI-korpuksessa 186 kertaa, UK-korpuksessa 300 kertaa ja EU-korpuksessa 684 kertaa. UK-korpuksessa sen sijaan *if* on käytetyin tarkastelluista ehtolauseilmauksista, mutta yleisesti vaihtelu on UK-korpuksessa suurempaa. FI-korpuksessa tarkasteltuja ehtolauseilmauksia on myös ylipäänsä käytetty huomattavasti enemmän kuin muissa korpuksissa. Yhteensä tarkasteltuja ehtolauseilmauksia esiintyy FI-korpuksessa 2 148 kertaa, UK-korpuksessa 1 648 kertaa ja EU-korpuksessa 1 778 kertaa.

Taulukko 4. Ehtolauseilmaukset korpuksissa

	FI-korpus	UK-korpus	EU-korpus	P-arvo	
				FI-korpus– UK-korpus	FI-korpus– EU-korpus
apart from	5	32	0	$P < 0,001$	$P = 0,04$
except	17	73	69	$P < 0,001$	$P < 0,001$
if	1 040	568	386	$P < 0,001$	$P < 0,001$
in so far as / insofar as	38	14	10	$P = 0,016$	$P < 0,001$
in the absence of	4	1	10	$P = 0,29$	$P = 0,049$
in (the) case(s)	56	181	157	$P < 0,001$	$P < 0,001$
in the event	4	2	35	$P = 0,62$	$P < 0,001$
notwithstanding	52	9	5	$P < 0,001$	$P < 0,001$
on (the) condition	26	0	4	$P < 0,001$	$P < 0,001$
provided that	34	3	60	$P < 0,001$	$P < 0,001$
save for	0	1	0	-	-
unless	148	105	58	$P = 0,53$	$P < 0,001$
when	386	94	204	$P < 0,001$	$P < 0,001$
whenever	3	15	9	$P < 0,001$	$P = 0,04$
where	186	300	684	$P < 0,001$	$P < 0,001$
who	149	250	87	$P < 0,001$	$P = 0,007$

Ehtolauseilmausten tarkastelu ei juuri tue tekstuaalista yhdenmukaisuutta korpusten välillä. Korpusten välillä on runsaasti vaihtelua sen suhteen, minkä verran ehtolauseilmauksia käytetään ja mitä ilmauksia ehtolauseissa käytetään. FI-korpuksen ehtolauseilmaukset ovat selkeästi yksinkertaistuneita verrattuna kahteen muuhun korpukseen, mutta FI-korpus ei ole ehtolauseilmauksiltaan selkeästi yhdenmukaisempi UK-korpuksen tai EU-korpuksen kanssa. Kuitenkin tilastollista merkitsevyyttä tarkasteltaessa esiin nousevat ilmaukset *in the absence of* sekä *unless*, joiden esiintymien taajuuden ero FI-korpuksen ja UK-korpuksen välillä on huomattavasti tilastollisesti merkitsemättömämpi kuin ero FI-korpuksen ja EU-korpuksen välillä, mikä saattaa kieliä samankaltaisuudesta FI-korpuksen ja UK-korpuksen välillä, sillä näiden korpusten erolla on kohtalainen todennäköisyys johtua satunnaisesta vaihtelusta. Ilmauksen *in the absence of* esiintyvyyden erolle FI-korpuksen ja UK-korpuksen välillä laskettu P-arvo on 0,29, kun taas erolle EU-korpukseen laskettu P-arvo on 0,049, joka on juuri ja juuri tilastollisesti melkein merkitsevä. Konjunktion *unless* erolle FI-korpuksen ja UK-korpuksen välillä laskettu P-arvo taas on 0,53, ja erolle EU-korpukseen laskettu P-arvo on alle 0,001, mikä kielii siitä, että konjunktion *unless* käyttö on samankaltaista FI-korpuksessa ja UK-korpuksessa.

Oikeuden, velvollisuuden ja luvan kommunikointiin liittyvinä ilmauksina tarkastellaan konkordanssihaun avulla verbejä *may*, *must*, *shall*, *permit*, *forbid*, *shall not*, *must not* sekä *may not*. Verbit *permit* ja *forbid* haetaan lemmahaualla, sillä tarkastelussa halutaan ottaa huomioon myös taivutetut muodot kuten *permitted* ja *forbidden*. Taulukossa 5 on esitetty tarkasteltujen oikeutta, velvollisuutta ja lupaa kommunikoivien performatiiviverbien esiintymät korpuksissa. Taulukon luvuista on poistettu kaksinkertaiset esiintymät (esim. *shall*-verbin esiintymät eivät sisällä *shall not* -ilmausta).

Korpuksat eroavat toisistaan huomattavasti oikeutta, velvollisuutta ja lupaa kommunikoivien verbien suhteen. *Shall*-performatiivi esiintyy 1 680 kertaa FI-korpuksessa, 929 kertaa UK-korpuksessa ja peräti 3 673 kertaa EU-korpuksessa. Täten *shall*-verbiä käytetään FI-korpuksessa 80,84 prosenttia enemmän kuin UK-korpuksessa ja 54,26 prosenttia vähemmän kuin EU-korpuksessa (P-arvo molempiin suuntiin alle 0,001). EU-korpuksessa *shallin* käyttö on siis huomattavasti taajempaa kuin muissa korpuksissa. *Must*-performatiivia sen sijaan käytetään EU-korpuksessa vain 35 kertaa, mikä on 92,72 prosenttia vähemmän kuin UK-korpuksessa (481 esiintymää) ja 93,57 prosenttia vähemmän kuin FI-korpuksessa (544 esiintymää). Myös UK-korpuksessa *shall* on yleisempi kuin *must*, vaikka *shallin* käyttöä pyritään välttämään (*Drafting Guidance*, 14). *Must*-performatiivia käytetään FI-korpuksessa ja

UK-korpuksessa lähestulkoon yhtä paljon: FI-korpuksessa sana esiintyy 13,1 prosenttia enemmän kuin UK-korpuksessa, ja eron P-arvoksi laskettiin tilastollisesti ei-merkittävä 0,57, kun taas erolle EU-korpukseen laskettu P-arvo on alle 0,001.

Taulukko 5. Oikeutta, velvollisuutta ja lupaa kommunikoivat performatiivit

	FI-korpus	UK-korpus	EU-korpus	P-arvo	
				FI-korpus–UK-korpus	FI-korpus–EU-korpus
may	1 044	1 501	697	P < 0,001	P < 0,001
must	544	481	35	P = 0,57	P < 0,001
shall	1 680	929	3 673	P < 0,001	P < 0,001
permit (verbi, lemma)	30	62	18	P < 0,001	P = 0,006
forbid (verbi, lemma)	4	0	0	P = 0,06	P = 0,02
shall not	98	121	199	P = 0,02	P < 0,001
must not	8	17	1	P = 0,04	P = 0,005
may not	88	82	13	P = 0,89	P < 0,001

Lupaa ilmaiseva performatiivi *may* on ylliedustettuna UK-korpuksessa kahteen muuhun korpukseen nähden; UK-korpuksessa *may* esiintyy 43,77 prosenttia enemmän kuin FI-korpuksessa ja 115,35 prosenttia enemmän kuin EU-korpuksessa, ja FI-korpuksen ero molempiin korpuksiin on tilastollisesti erittäin merkittävä. Sen sijaan kiellon *may not* esiintyy FI-korpuksessa vain 7,3 prosenttia enemmän kuin UK-korpuksessa, ja eron P-arvo on 0,89, kun taas erolle EU-korpukseen laskettu P-arvo on tilastollisesti erittäin merkittävä alle 0,001. Myös kiellon *shall not* suhteen FI-korpus on lähempänä UK-korpusta kuin EU-korpusta: FI-korpuksen ja UK-korpuksen välinen ero on tilastollisesti melkein merkittävä (P = 0,02), kun taas ero EU-korpukseen on erittäin merkittävä (P < 0,001).

Muiden tarkasteltujen ilmausten esiintymien lukumäärät ovat niin pieniä tai korpuksen välillä niin poikkeavia, ettei niistä voi tehdä johtopäätöksiä. On kuitenkin ainakin selvää, että verbiä *forbid* sekä kiellon *must not* ei juuri käytetä minkään korpuksen teksteissä.

Oikeuksien, velvollisuuksien, luvan ja kiellon ilmauksien suhteen FI-korpus näyttäisi olevan lähempänä UK-korpusta kuin EU-korpusta. Yksinkertaistumishypoteesiin nojaten oikeuden

logiikan kommunikoinnin tulisi olla käännöksissä yksinkertaistunutta. Toisin sanoen yksittäisten ehtolauseilmausten sekä oikeutta, velvollisuutta ja lupaa kommunikoivien ilmausten tulisi olla ylliedustuneita FI-korpuksessa EU-korpukseen ja UK-korpukseen nähden, ja näissä alun perin kohdekielisissä korpuksissa puolestaan pitäisi olla enemmän variaatiota. Tästä ei havaittu merkkejä oikeutta, velvollisuutta ja lupaa kommunikoivien ilmausten osalta, mutta ehtolauseilmauksista *if*-konjunktio on ylliedustunut FI-korpuksessa, mikä saattaa kieliä yksinkertaistumisesta ehtolauseissa.

4.3 Analyysin yhteenveto

Tässä tutkimuksessa tarkasteltiin, ovatko suomalaisten lakien epäviralliset englanninkieliset käännökset tekstuaalisilta piirteiltään yhdenmukaisempia Euroopan unionin asetusten englanninkielisten toisintojen kanssa vai Yhdistyneen kuningaskunnan parlamentin lakien kanssa. Tutkimuksessa pyrittiin myös vastaamaan siihen, pätevätkö käännösuniversaalihypoteesit lainsäädännön käännöksissä. Tutkimusaineiston korpusanalyysissa saatiin selville, että suomalaisten lakien käännöksistä koostuva FI-korpus on tekstuaalisilta piirteiltään lähempänä Euroopan unionin asetuksista koostuvaa EU-korpusta kuin Yhdistyneen kuningaskunnan parlamentin laeista koostuvaa UK-korpusta. Tarkastelluille käännösuniversaalihypoteeseille sen sijaan aineiston analyysistä löytyi tukea vain osittain.

4.3.1 Korpusten tekstuaalinen yhdenmukaisuus

Tutkimusaineiston analyysissa havaittiin, että suomalaisten lakien englanninkielisistä käännöksistä koostuva korpus (*FI-korpus*) on tekstuaalisilta piirteiltään lähempänä Euroopan unionin asetusten englanninkielisistä toisinoista koostuvaa korpusta (*EU-korpus*) kuin Yhdistyneen kuningaskunnan parlamentin antamista laeista koostuvaa korpusta (*UK-korpus*). Tuloksissa havaittiin kuitenkin vaihtelua makro- ja mikrotason välillä.

Makrotasolla FI-korpus oli johdonmukaisesti tekstuaalisesti yhdenmukaisempi EU-korpuksen kuin UK-korpuksen kanssa. FI-korpuksen keskimääräinen virkepituus, sanaluokkajakauma, leksikaalinen tiheys ja TTR-luku seurasivat melko tarkoin EU-korpuksen arvoja, kun taas FI-korpuksen ja UK-korpuksen väliset erot olivat huomattavasti suurempia. Lisäksi arvojen keskihajonta virkepituuksissa ja standardisoiduissa TTR-luvuissa oli samansuuntaista EU-korpuksen kanssa UK-korpuksen vaihtelun ollessa suurempaa.

Mikrotasolla havaittiin enemmän eroavaisuuksia myös FI-korpuksen ja EU-korpuksen välillä, vaikka makrotasolla nämä kaksi korpusta olivat keskenään hyvin samankaltaisia. Sana- ja

n-grammifrekvenssilistoissa vaihtelu oli suurempaa, ja FI-korpus näytti tekstuaalisia samankaltaisuuksia ja eroavaisuuksia sekä EU-korpukseen että UK-korpukseen nähden mutta oli kuitenkin jokseenkin yhdenmukaisempi EU-korpuksen kanssa. Juridisen logiikan kommunikointiin liittyvän sanaston suhteen FI-korpus ei näyttäytynyt yhdenmukaisempaan kumpaankaan alun perin kohdekieliseen korpukseen nähden.

4.3.2 Käännösuniversaalihypoteesit

Tutkimuksessa käännösuniversaalihypoteeseille ei löydetty johdonmukaisesti tukea, mutta tulokset eivät olleet myöskään yksinomaan ristiriidassa käännösuniversaalihypoteesien kanssa. Makrotasolla FI-korpuksessa ei havaittu merkkejä eksplikoitumisesta. Viitteitä yksinkertaistumisesta havaittiin sanaluokkien analyysissä, mutta leksikaalisen tiheyden ja TTR-luvun analyysissä saadut tulokset olivat ristiriidassa yksinkertaistumishypoteesin kanssa. Normalisoitumista havaittiin virkepituuksien keskihajonnassa sekä tuhannen sanan välein lasketussa standardisoidussa TTR-luvussa. Normalisoituminen oli huomattavampaa UK-korpukseen nähden, ja vaikka FI-korpus oli normalisoitunut myös EU-korpukseen verrattuna, oli tämän normalisoitumisen aste melko pientä ja tilastollisesti merkitsemätöntä. Mikrotasollakaan ei löydetty merkkejä eksplikoitumisesta. Yksinkertaistumisesta havaittiin joitakin merkkejä juridista logiikkaa kommunikoivissa ilmauksissa sekä sana- ja n-grammilistoissa. Normalisoitumista ei tarkasteltu mikrotasolla.

Käännösuniversaalihypoteeseja tukevat tulokset saatiin lähinnä verratessa FI-korpusta UK-korpukseen, ja vaikka EU-korpukseenkin verrattuna saatiin samoja tuloksia, oli EU-korpus tekstuaalisilta piirteiltä lähes yksinomaan lähempänä FI-korpusta kuin UK-korpus. Mahdollinen normalisoituminen ja yksinkertaistuminen FI-korpuksen ja EU-korpuksen välillä oli siis hyvin pientä, ja erot korpusten välillä olivat enimmäkseen tilastollisesti merkitsemättömiä, joskin juridista logiikkaa kommunikoivissa ehtolauseissa huomattiin selvää yksinkertaistumista myös EU-korpukseen nähden.

5 Lopuksi

Tässä tutkielmassa tarkasteltiin suomalaisen lainsäädännön epävirallisten käännösten englantia suhteessa Euroopan unionin asetusten englanninkielisiin toisintoihin sekä Yhdistyneen kuningaskunnan parlamentin antamiin lakeihin. Tavoitteena oli saavuttaa näkemys siitä, palautuuko suomalaisten juridisten tekstien käännösenglanti tekstuaalisesti pikemmin juridiseen euroopanenglantiin vai common law -perinteeseen.

Ottaen huomioon, että suomalainen oikeusjärjestelmä pohjautuu mannereurooppalaiseen oikeusperinteeseen, jota common law -oikeuden englanti ei ole muotoutunut kuvaamaan, oli oletettavaa, että käännöskorpus on lähempänä Euroopan unionin asetusten korpusta. Euroopan unionin instituutioissa muotoutunutta juridista euroopanenglantia on kuitenkin käsitelty oikeuslingvistiksi ja käännöstieteellisessä kirjallisuudessa verrattain vähän ja useimmiten terminologisesta näkökulmasta, joten tutkimuksen tekstuaalinen näkökulma tarjosi vahvistusta sille, että eurooppalainen juridinen *lingua franca* -englanti poikkeaa common law -englannista huomattavasti myös tekstuaalisesti.

Tutkimuksessa suomalaisten lakien käännösten korpuksen havaittiin olevan tekstuaalisilta piirteiltään yhdenmukaisempi Euroopan unionin asetuksista koostuvan korpuksen kanssa kuin Yhdistyneen kuningaskunnan laeista koostuvan korpuksen kanssa. Korpusaineistosta saadut tulokset olivat melko johdonmukaisia tarkasteltaessa korpuksia globaalisti makrotasolla, mutta mikrotasolla saaduissa havainnoissa oli enemmän vaihtelua, eivätkä suomalaisten lakien käännökset näyttäneet mikrotasolla olevan sinänsä juuri lähempänä Euroopan unionin asetuksia tai Yhdistyneen kuningaskunnan lakeja. Makrotason analyysin perusteella voi kuitenkin todeta, että virkerakenteiltaan suomalaisen lainsäädännön käännösenglanti on selvästi lähempänä EU-englantia kuin common law -oikeusjärjestelmien englantia. Tuloksista kävi myös ilmi, että Euroopan unionin lainsäädännön englanti on common law -oikeuden englantiin nähden selvästi erilaista tekstuaalisilta rakenteiltaan.

Korpusaineiston analyysin apuna käytettiin eksplikoitumisen, normalisoitumisen ja yksinkertaistumisen käännösuniversaalihypoteeseja. Siinä missä käännösuniversaalihypoteesit toimivat hyvin työkaluna korpusten tekstuaalisten piirteiden vertailuun, ei eksplikoitumisen ja yksinkertaistumisen hypoteeseille sinänsä löydetty korpusaineistosta johdonmukaista tukea. Normalisoitumista aineistossa sen sijaan havaittiin virkepituuksien ja standardisoidun TTR-luvun osalta, mutta nämäkin erot olivat pieniä ja erityisesti FI-korpuksen ja EU-korpuksen

välillä tilastollisesti merkitsemättömiä. Vaikka käännöskorpuksen aineisto koostui nimenomaan lainsäädännön epävirallisista käännöksistä, joiden laatimisen voisi kuvitella käännöksiä todistusvoimaisuuden puutteesta johtuen olevan vapaampaa kuin esimerkiksi suomalaisten lakien virallisten ruotsinkielisten käännösten kohdalla, ei käännösten universaaliksi ehdotettuja piirteitä havaittu siten kuin on ehdotettu. Mikäli käännösaineisto olisi koottu todistusvoimaisista juridisista teksteistä, kuten esimerkiksi sopimuskäännöksistä, käännösten suhde käännösuniversaalihypoteeseihin saattaisi olla vielä ristiriitaisempi. Käännösuniversaalihypoteesien suhteen saadut tulokset kielivät yleisesti ottaen siitä, että lainsäädännön käännökset poikkeavat luonteeltaan hyvin paljon muusta kääntämisestä. Eri oikeusjärjestelmien välillä luoviminen ja terminologisen vastaamattomuuden kompensointi, juridisilta käännöksiltä vaadittava uskollisuus lähdetekstin muotoiluille sekä institutionaaliset tekstien laatimisen käytännöt ovat omiaan aiheuttamaan eroja muuhun erikoisalakääntämiseen nähden, kaunokirjallisten tekstien kääntämisestä puhumattakaan.

Se, että käännöskorpus koostui epävirallisista säädöskäännöksistä, ei sinänsä näkynyt tekstuaalisten piirteiden yhdenmukaisuutta makrotasolla tarkasteltaessa, mutta mikrotason ilmiöiden havainnointiin sillä saattoi olla merkitystä, että käännöksiä ei välttämättä ole tehty vaadittavilla juridisten tekstien kääntämisen resursseilla ja kompetensseilla. Epäviralliset säädöskäännökset myös laadittaneen useimmiten lainsäädäntöprosessista erillään toisin kuin esimerkiksi suomalaisten lakien ruotsinkieliset käännökset, millä voi olla vaikutusta esimerkiksi juridisten merkityssuhteiden kommunikointiin käännöksissä. Lisäksi mikäli Kinnusen (2006, 21) väite siitä, että vieraskielisiä informatiivisia käännöksiä lukevien tuomareiden oletetaan tuntevan vieraan oikeusjärjestelmän peruskäsitteet, pitää paikkansa, ei epävirallisissa käännöksissä liene sen takia juurikaan käytetty selittäviin parafraaseihin viittaavia ilmauksia, kuten tarkennetta *which*.

Korpusaineistoon eri lainsäädäntöportaaleista poimittu otos oli kuitenkin siinä mielessä edustava, että tutkimuksessa saadut tulokset olivat johdonmukaisia makrotasolla, jolla FI-korpuksen ja EU-korpuksen havaittiin olevan selkeästi keskenään tekstuaalisesti yhdenmukaisempia kuin FI-korpuksen ja UK-korpuksen. Toisaalta vielä isommasta otoksesta olisi saattanut olla apua mikrotason ilmiöiden tutkimisessa. Tähän tutkimukseen kerätty otos oli melko maltillisen kokoinen, mutta isommassa tutkimusprojektissa mahdollista olisi tietysti sisällyttää korpukseen kaikki Finlex-lainsäädäntöportaalissa julkaistut englanninkieliset säädöskäännökset.

Tutkimus ei ollut tyhjentävä myöskään siksi, että tarkastellut Yhdistyneen kuningaskunnan parlamentin antamat lait ovat ainoastaan yksi esimerkki common law -oikeusjärjestelmän juridisesta englannista. Korpustutkimusta common law’sta tarvitaan siis lisää, ja vertailuun tulee sisällyttää laajempi kirjo eri instituutioissa laadittuja tekstejä eri englanninkielisistä maista, kuten Irlannista, Skotlannista, Maltalta ja Australiasta, joista Irlanti ja Skotlanti käsitetään hybridijärjestelmiksi, Malta mannereurooppalaiseen oikeusperheeseen kuuluvaksi oikeusjärjestelmäksi ja Australia common law -oikeusjärjestelmäksi.

Vertailevan korpustutkimuksen menetelmä sopi tutkimuksen tarkoituksiin hyvin ja menetelmällä saadut tulokset tarjosivat tuoretta tietoa suomalaisen juridisen käännösendlannin luonteesta ja suhteutumisesta juridiseen euroopanenglantiin ja common law -englantiin. Toisaalta tutkimus kaipasi tuekseen vielä kvalitatiivista tekstianalyysia. Korpustyökalulla kyetään havainnoimaan, että suomalaisen lainsäädännön englanninkieliset käännökset ovat tekstuaalisilta piirteiltään lähempänä juridista euroopanenglantia kuin common law -englantia, mutta syitä on vaikea hahmottaa, sillä esimerkiksi institutionaaliset käytännöt voivat vaikuttaa olennaisesti myös kielellisiin piirteisiin. Syitä tutkimustulosten taustalla pitäisikin etsiä kvalitatiivisen tekstianalyysin menetelmin, mihin tämän tutkimuksen puitteissa ei ollut mahdollisuutta.

Mielenkiintoista olisi tietysti myös suomalaisten lakien suomenkielisten lähdetekstien vertailu englanninkielisiin kohdeteksteihin paralleelikorpuksen avulla. Tällä tavoin voitaisiin havainnoida paremmin suomalaisten lakien englanninkielisissä käännöksissä käytettyjä käännösstrategioita. Käännösuniversaaliteorian näkökulmasta paralleelikorpuksen avulla olisi mahdollista myös tarkastella lainsäädännön käännösten suhteutumista ehdotettuihin S-käännösuniversaaleihin ja tutkia, millaisina juridiset käännökset näyttäytyvät lähde-kohde-verailussa muiden tekstilajien käännöksiin verrattuna. Paralleelikorpuksesta myös eksplikoitumista olisi kenties mielekkäämpiä tutkia, sillä tässä tutkimuksessa eksplikoitumisen käsittely jäi pintapuoliseksi eikä merkittäviä tuloksia eksplikoitumisesta saavutettu. On otettava myös huomioon, että tämän tutkimuksen käännöskorpus koostui epävirallisista lainsäädännön käännöksistä, minkä vuoksi eksplisiittisiä selityksiä lähdetekstin termeille ei ollut ehkä käytetty niin laajalti kuin oletettiin. Eksplikoitumisesta ei löydetty viitteitä esimerkiksi selittävien parafrasien osalta, mutta olisi mielenkiintoista verrata esimerkiksi virallisia sopimuskäännöksiä alun perin kohdekielellä laadittuihin sopimuksiin ja selvittää kvalitatiivisin menetelmin, onko eksplikoituminen näkyvämpää juridisten tekstien virallisissa ja

todistusvoimaisissa käännöksissä. Yksi vaihtoehto olisi myös tutkia suomalaisten säädösten virallisia ruotsinkielisiä käännöksiä ruotsalaisiin säädöksiin, joskin ruotsalaisen ja suomalaisen oikeusjärjestyksen välillä ei välttämättä terminologista vastaamattomuutta juuri ilmene johtuen maiden yhteisestä historiasta ja jaetusta pohjoismaisesta oikeusperheestä.

Ottaen huomioon, miten riippuvaisia lainsäädännölliset tekstit ja niiden käännökset ovat niistä instituutioista, joissa niitä laaditaan, tutkimuksen teoreettisessa viitekehyksessä olisi kenties ollut syytä kartoittaa tarkemmin institutionaalisuutta ja sen vaikutuksia kääntämiseen. Havaittuun juridisen euroopanenglannin käännöksenomaisuuteen vaikuttaa epäilemättä se, millaisia käytäntöjä Euroopan unionin instituutioissa on tekstien laatimiselle. Juridiset tekstit, erityisesti lainsäädännölliset tekstit, ovatkin kenties muita tekstilajeja riippuvaisempia institutionaalisista tekstin laatimisen käytännöistä. Institutionaalisuudella on todennäköisesti vaikutusta myös tutkimuksessa havainnoituun lainsäädännön käännösten osittaiseen yhteensopimattomuuteen käännösuniversaalihypoteesien kanssa. Vaikka suomalaisten lakien englanninkieliset käännökset ovat epävirallisia, on valtioneuvoston kanslia antanut tarkat ohjeistukset myös suomalaisten säädösten kääntämiseen englanniksi (ks. Säädösten kääntäminen englanniksi ja valtiosopimusten suomentaminen). Täten institutionaalisuus on läsnä myös epävirallisissa lainsäädännön käännöksissä.

Euroopan unionin lainsäädännöllisten tekstien englanti näyttäytyi siis tässä tutkimuksessa tekstuaalisesti samankaltaisena kuin suomalaisen lainsäädännön käännösten englanti. Vaikka valtaosa Euroopan unionissa laadittavista lainsäädännöllisistä teksteistä laaditaan alun perin englanniksi, tekstejä laativat unionin instituutioissa usein ei-natiivipuhujat (ks. luvut 2.1.4 & 2.2.2). Tekstit myös laaditaan alun alkujaan käännettäviksi: Euroopan unionin lainsäädäntö julkaistaan kaikilla 24 virallisella kielellä, joilla kaikki lainsäädäntö on todistusvoimaista. Euroopan unionin monikielisyys asettaa vaatimuksia sille, miten pitkiä esimerkiksi virkkeet voivat olla ja miten ne voivat rakentua; Yhdistyneen kuningaskunnan lainsäädännöllisissä teksteissä tällaisia rajoituksia ei ole. Sinänsä ei ole siis yllättävää, että juridinen euroopanenglanti muistuttaa tekstuaalisesti suomalaisten juridisten tekstien käännösendlantia.

Siihen nähden, että englanti on 1990- ja 2000-luvuilla syrjäyttänyt ranskan Euroopan unionin valtakielenä, ei Euroopan unionissa käytettävän juridisen englannin tekstuaalisista piirteistä tiedetä kuitenkaan vielä juuri mitään. Toisaalta oikeuslingvistikseen näkökulmasta common law -englannin ja juridisen euroopanenglannin olennaisimmat erot liittynevät terminologisiin kysymyksiin. Termityö Euroopan unionissa on ottanut harppauksen 1990-luvun lopulta lähtien

EU:n toimielinten yhteisen käännös- ja tulkkauskomitean perustamisen myötä (Mattila 2017, 188–189); komitean tuotoksia on muun muassa unionin julkinen termipankki IATE (mp.). Mattilan mukaan termien käyttöönotto satunnaisesti ja epäjohdonmukaisesti on vaarana monimutkaisessa käännösorganisaatiossa (mp.), joten voi hyvin olla, että prioriteettina syntyneessä olevan juridisen englannin variantin käytäntöjen harmonisoinnissa on yhtenäisen ja tarkan termistön luominen, jonka jälkeen voidaan vasta keskittyä tekstuaalisten piirteiden yhtenäistämiseen. Myös tämän vuoksi EU-englanti saattaa näyttäytyä korpusanalyysissä makrotasolla tarkasteltuna käännöksenomaisena. Termistölähtöinen tutkimusote olisi voinut soveltua tässäkin tutkimuksessa paremmin mikrotason ilmiöiden tarkasteluun, mutta tutkimuksessa saatiin kuitenkin arvokasta tietoa juridisen euroopanenglannin tekstuaalisiin piirteisiin liittyen, minkä pohjalta tarkempi kvalitatiivinen tutkimus voi olla tarpeen.

Lähdeluettelo

- Act on Child Custody and Right of Access*. Oikeusministeriö. Saatavilla: https://www.finlex.fi/en/laki/kaannokset/1983/en19830361_20090436.pdf. [Luettu 29.5.2019.]
- Andström, Kristina 2017. Lainoppi ja oikeudenalat. Teoksessa Kristina Andström, Paula Klami-Wetterstein, Ilona Nieminen, Mia Oksala, Anni Tuomela & Tero Virtakainen, *Johdatus oikeustieteeseen*. [Verkkomateriaali]. Saatavilla: <https://blogs.helsinki.fi/avoin-johdatusoik/lainoppi-ja-oikeudenalat/> [Luettu 15.3.2019.]
- Ángeles, María & Orts Llopis 2017. Legal English as the Lingua Franca for International Law. Traits and Pitfalls for ESP Practitioners and Legal Translators. Teoksessa Nicole Keng, Anita Nuopponen & Daniel Rellstab (toim.), *Ääniä, Röster, Voices, Stimmen. VAKKI-symposiumi XXXVII 9. 10.2.2017. Vakki Publications 8*. 17–28. Saatavilla: http://www.vakki.net/publications/2017/VAKKI2017_Orts_Llopis.pdf [Luettu 1.5.2019.]
- Baaij, Cornelis J.W. 2012. The EU Policy on Institutional Multilingualism: Between Principles and Practicality. *International Journal of Language and Law* 1. 14–32. Saatavilla: <https://www.languageandlaw.eu/jll/article/view/4> [Luettu 20.5.2019.]
- Bajčić, Martina 2018. The Role of EU Legal English in Shaping EU Legal Culture. *International Journal of Language & Law* 7. 8–24. Saatavilla: <https://www.languageandlaw.eu/jll/article/view/48> [Luettu 22.4.2019.]
- Baker, Mona 1993. Corpus Linguistics and Translation Studies. Implications and Applications. Teoksessa Mona Baker, Gill Francis & Elena Tognini-Bonelli (toim.), *Text and Technology. In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins. 233–250.
- Bhatia, Vijay K. 2004. *Worlds of Written Discourse*. London: Continuum.
- Biel, Łucja 2014a. *Lost in the Eurofog: The textual fit of translated law*. Frankfurt am Main: Peter Lang.
- Biel, Łucja 2014b. Phraseology in Legal Translation: A Corpus-Based Analysis of Textual Mapping in EU Law. Teoksessa Le Cheng, King Kui Sin & Anne Wagner (toim.), *The Ashgate Handbook of Legal Translation*. Farnham: Ashgate Publishing. 177–192.
- Biel, Łucja 2015. Phraseological profiles of legislative genres: complex prepositions as a special case of legal phrasemes in EU law and national law. *Fachsprache* 37 (3–4). 139–160. Saatavilla: <https://doi.org/10.24989/fs.v37i3-4.1286> [Luettu 21.5.2019.]
- Bogdan, Michael 1993. *Komparativ rättskunskap*. 2. painos. Tukholma: Norstedts juridik.
- Chesterman, Andrew 2004. Hypotheses about translation universals. Teoksessa Gyde Hansen, Kirsten Malmkjær & Daniel Gile (toim.), *Claims, Changes and Challenges in Translation Studies. Selected contributions from the EST Congress, Copenhagen 2001*. Amsterdam/Philadelphia: John Benjamins. 1–13.
- Chesterman, Andrew 2010. Translation universals. Teoksessa Yves Gambier & Luc van Doorslaer (toim.), *Handbook of Translation Studies, Volume 2*. Philadelphia: John Benjamins Publishing Company. 175–179.

- Doczekalska, Agnieszka 2018. Comparing Multilingual Practices in the EU and the Canadian Legal Systems: The Case of Terminological Choices in Legislative Drafting. Teoksessa Prieto Ramos (toim.). 102–112.
- Drafting Guidance*. Office of the Parliamentary Counsel. Saatavilla: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/61010/Office_of_the_Parliamentary_Counsel_revised_guidance_16_12_11.pdf [Luettu 15.3.2019.]
- Euroopan unionin toimintaperiaatteet*. [Verkkodokumentti]. Euroopan parlamentin Suomen tiedotustoimisto. Saatavilla: <https://europarlamenti.info/fi/Euroopan-unioni/toimintaperiaatteet/> [Luettu 6.1.2019.]
- Euroopan unionista tehdyn sopimuksen konsolidoitu toisinto*. Euroopan unionin virallinen lehti 26.10.2010. Saatavilla: https://eur-lex.europa.eu/resource.html?uri=cellar:2bf140bf-a3f8-4ab2-b506-fd71826e6da6.0004.02/DOC_1&format=PDF [Luettu 29.5.2019.]
- Felici, Annarita 2015. Translating EU Legislation from a Lingua Franca: Advantages and Disadvantages. Teoksessa Susan Šarčević (toim.), *Language and Culture in EU Law : Multidisciplinary Perspectives*. London: Routledge. 123–139.
- House, Juliane 2008. Beyond Intervention: Universals in Translation? *Trans-kom* 1(1), 6–19. Saatavilla: http://www.trans-kom.eu/bd01nr01/trans-kom_01_01_02_House_Beyond_Intervention.20080707.pdf [Luettu 21.5.2019.]
- Husa, Jaakko 2013. *Oikeusvertailu*. Helsinki: Lakimiesliiton Kustannus.
- Jäsenvaltion oikeus – Englanti ja Wales*. [Verkkodokumentti]. Euroopan oikeusportaali. Saatavilla: https://e-justice.europa.eu/content_member_state_law-6-ew-fi.do?init=true&member=1 [Luettu 10.5.2019.]
- Jäsenvaltion oikeus – Suomi*. [Verkkodokumentti]. Euroopan oikeusportaali. Saatavilla: https://e-justice.europa.eu/content_member_state_law-6-fi-fi.do?member=1 [Luettu 10.5.2019.]
- Jokela, Antti 2005. *Oikeudenkäynnin perusteet*. 2. painos. Helsinki: Talentum Oyj.
- Kielilaki 423/2003*. Annettu Helsingissä 6.6.2003. Saatavilla: <https://www.finlex.fi/fi/laki/alkup/2003/20030423> [Luettu 29.5.2019.]
- Kinnunen, Tuija 2006. *Käännös oikeudenkäyntiaineistona : Lakitekstien tulkinta oikeuskäsitteiden kääntämisen näkökulmasta*. Tampereen yliopisto, väitöskirja.
- Krogsgaard Vesterager, Anja 2017. Explicitation in legal translation – a study of Spanish-into-Danish translation of judgments. *The Journal of Specialised Translation* 27. 104–123. Saatavilla: https://www.jostrans.org/issue27/art_krogsgaard.pdf [Luettu 21.5.2019.]
- Mauranen, Anna 2004. Corpora, universals and interference. Teoksessa Mauranen & Kujamäki (toim.). 65–82.
- Lainvalmistelun prosessiopas*. [Verkkodokumentti]. Finlex. Saatavilla: <http://lainvalmistelu.finlex.fi/> [Luettu 13.4.2019.]
- Laviosa, Sara 2003. Corpus and Simplification in Translation. Teoksessa Susan Perilli (toim.), *Translation Translation*. Amsterdam: Rodopi. 153–162.

- Kjær, Anne Lise 2000. On the structure of legal knowledge: The importance of knowing legal rules for understanding legal texts. Teoksessa Lita Lundquist & Robert J. Jarvella (toim.), *Language, text, and knowledge : mental models of expert communication*. Berlin: Mouton de Gruyter.
- Mattila, Heikki E. S. 2008. Euroopan oikeudellinen integraatio ja käännösongelmat. Teoksessa Markus Aarto & Markku Vartiainen (toim.), *Oikeus kansainvälisessä maailmassa : Ilkka Saraviidan juhlakirja*. Helsinki: Edita. 147–160.
- Mattila, Heikki E. S. 2017. *Vertaileva oikeuslingvistiikka. Juridinen kielenkäyttö, lakimieslatina, kansainväliset oikeuskielet*. 2. painos. Helsinki: Alma Talent.
- Mauranen, Anna 2006. Translation universals. Teoksessa Keith Brown (toim.), *Encyclopedia of Language and Linguistics. Vol. 13*. Oxford: Elsevier. 93–100.
- Mauranen, Anna & Pekka Kujamäki (toim.) 2004. *Translation Universals: Do they Exist?* Amsterdam: John Benjamins.
- Nordman, Liselott 2009. *Lagöversättning som process och produkt*. Helsingin yliopisto, väitöskirja.
- Ojala, Elina 2015. *Kieli, kääntäminen ja instituutio – onko institutionaalisten tekstien kieli yksinkertaisempaa käännöksissä?* Tampereen yliopisto, pro gradu -tutkielma.
- Paunio, Elina 2011. *Beyond Words – the European Court of Justice and Legal Certainty in Multilingual EU Law*. Helsingin yliopisto, väitöskirja.
- Piehl, Aino 2003: ”Selkeyttä vain” – EU-kieleen. *Kielikello* 4/2003. Saatavilla: <https://www.kielikello.fi/-/selkeytta-vain-eu-kieleen> [Luettu 18.3.2019.]
- Piehl, Aino 2006. The Influence of EU Legislation on Finnish Legal Discourse. Teoksessa Maurizio Gotti & Davide S. Giannoni (toim.), *New Trends in Specialized Discourse Analysis*. Bern: Peter Lang. 183–194.
- Prieto Ramos, Fernando (toim.) 2018. *Institutional Translation for International Governance. Enhancing Quality in Multilingual Legal Communication*. London: Bloomsbury Academic.
- Puurtinen, Tiina 2004. Explication of causal relations. Teoksessa Mauranen & Kujamäki (toim.). 165–176.
- Robertson, Colin 2012. EU Legal English: Common Law, Civil Law, or a New Genre? *European Review of Private Law* 5–6. 1215–1240.
- Sajama, Seppo 2017. *Oikeudellinen argumentaatio 2017*. [Kurssimateriaali]. Saatavilla: <https://www.uef.fi/documents/300201/565073/Oikeudellinen+argumentaatio2017/6644e504-f3be-469d-a33c-d42ba92486c9> [Luettu 20.3.2019.]
- Samuel, Geoffrey 1994. *The Foundations of Legal Reasoning*. Tilburg: Maklu.
- Šarčević, Susan 1997. *New Approach to Legal Translation*. The Hague: Kluwer Law International.
- Šarčević, Susan 2018. Challenges to Legal Translators in Institutional Settings. Teoksessa Prieto Ramos (toim.). 9–24.
- Schedules*. [Verkkodokumentti]. Yhdistyneen kuningaskunnan parlamentti. Saatavilla: <https://www.parliament.uk/site-information/glossary/schedules/> [Luettu 28.11.2018.]
- Säädösten kääntäminen englanniksi ja valtiosopimusten suomentaminen*. Valtioneuvoston kanslia. Saatavilla: <http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/79776/>

[J09_2017_Saadosten%20kaantaminen%20englanniksi_KANSILLA_FI.pdf?sequence=1&isAllowed=y](#) [Luettu 8.5.2019.]

Vuorinen, Erkki 2017. *Kohti vuotta 2019 – millaiset ovat suomen näkymät seuraavan puheenjohtajuuskauden lähestyessä?* [Luentomateriaali]. Saatavilla: <https://www.sktl.fi/liitto/suomi-kaantaa-100-vuotta-kaannet/eu-ta-suomeksi-ja-eu-suomea-kaan/> [Luettu 26.11.2018.]

Williams, Christopher 2014. The future of ESP studies: building on success, exploring new paths, avoiding pitfalls. *ASp* 66. Saatavilla: <https://journals.openedition.org/asp/4616> [Luettu 2.5.2019.]

Liite 1: Lista aineistoon poimituista säädöksistä

Suomalaisten lakien käännökset (FI-korpus)

- 1) Act on Compensation for Environmental Damage (Haettu 8.11.2018.)
- 2) Act on Consideration for the Energy and Environmental Impact of Vehicles in Public Procurement (Haettu 8.11.2018.)
- 3) Act on Environmental Impact Assessment Procedure (Haettu 8.11.2018.)
- 4) Act on Environmental Permit Authorities (Haettu 8.11.2018.)
- 5) Act on Environmental Protection in Maritime Transport (Haettu 8.11.2018.)
- 6) Act on Implementation of the Legislation on Environmental Protection (Haettu 8.11.2018.)
- 7) Act on Jointly Owned Forests (Haettu 8.11.2018.)
- 8) Act on Restriction of the Use of Certain Hazardous Substances in Electrical and Electronic Equipment (Haettu 8.11.2018.)
- 9) Act on the Financing of Sustainable Forestry (Haettu 8.11.2018.)
- 10) Act on the Forestry Centres and Forestry Development Centre Tapio (Haettu 8.11.2018.)
- 11) Act on the Remediation of Certain Environmental Damages (Haettu 8.11.2018.)
- 12) Act on Trade in Forest Reproductive Material (Haettu 8.11.2018.)
- 13) Act on Verification and Notification of Origin of Electricity (Haettu 8.11.2018.)
- 14) Climate Change Act (Haettu 8.11.2018.)
- 15) Electrical Safety Act (Haettu 8.11.2018.)
- 16) Electricity Market Act (Haettu 8.11.2018.)
- 17) Environmental Damage Insurance Act (Haettu 8.11.2018.)
- 18) Environmental Protection Act (Haettu 8.11.2018.)
- 19) Forest Act (Haettu 8.11.2018.)
- 20) Forest Damages Prevention Act (Haettu 8.11.2018.)
- 21) Forest Management Association Act (Haettu 8.11.2018.)
- 22) Gene Technology Act (Haettu 8.11.2018.)
- 23) Nature Conservation Act (Haettu 8.11.2018.)
- 24) Temporary Act on the Financing of Sustainable Forestry (Haettu 8.11.2018.)
- 25) Waste Act (Haettu 8.11.2018.)
- 26) Waste Tax Act (Haettu 8.11.2018.)

Yhteensä 198 955 sanetta

Yhdistyneen kuningaskunnan parlamentin lait (UK-korpus)

- 1) Antarctic Act 2013 (Haettu 12.11.2018.)
- 2) Automated and Electric Vehicles Act 2018 (Haettu 12.11.2018.)
- 3) Climate Change and Sustainable Energy Act 2006 (Haettu 8.11.2018.)
- 4) Electricity (Miscellaneous Provisions) Act 2003 (Haettu 8.11.2018.)
- 5) Energy Act 2013 (Haettu 8.11.2018.)
- 6) Environment Act 1995 (Haettu 8.11.2018.)
- 7) Flood and Water Management Act 2010 (Haettu 12.11.2018.)
- 8) Gas Act 1995 (Haettu 12.11.2018.)
- 9) Green Energy (Definition and Promotion) Act 2009 (Haettu 12.11.2018.)
- 10) Home Energy Conservation Act 1995 (Haettu 12.11.2018.)
- 11) Land Drainage Act 1994 (Haettu 12.11.2018.)
- 12) Natural Environment and Rural Communities Act 2006 (Haettu 8.11.2018.)
- 13) Nuclear Safeguards Act 2018 (Haettu 12.11.2018.)
- 14) Plant Varieties Act 1997 (Haettu 12.11.2018.)
- 15) Sustainable Communities Act 2007 (Haettu 12.11.2018.)
- 16) Waste and Emissions Trading Act 2003 (Haettu 12.11.2018.)
- 17) Waste Minimisation Act 1998 (Haettu 12.11.2018.)
- 18) Water Industry Act 1999 (Haettu 12.11.2018.)

Yhteensä 196 368 sanetta

Euroopan unionin asetukset (EU-korpus)

- 1) Regulation (EU) No 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species (Haettu 12.11.2018.)
- 2) Commission Regulation (EC) No 1737/2006 of 7 November 2006 laying down detailed rules for the implementation of Regulation (EC) No 2152/2003 of the European Parliament and of the Council concerning monitoring of forests and environmental interactions in the Community (Haettu 21.11.2018.)
- 3) Commission Regulation (EU) 2017/2195 of 23 November 2017 establishing a guideline on electricity balancing (Haettu 21.11.2018.)

- 4) Commission Regulation (EU) No 1031/2010 of 12 November 2010 on the timing, administration and other aspects of auctioning of greenhouse gas emission allowances pursuant to Directive 2003/87/EC of the European Parliament and of the Council establishing a scheme for greenhouse gas emission allowances trading within the Community (Haettu 21.11.2018.)
- 5) Commission Regulation (EU) No 312/2014 of 26 March 2014 establishing a Network Code on Gas Balancing of Transmission Networks (Haettu 21.11.2018.)
- 6) Commission Regulation (Euratom) No 66/2006 of 16 January 2006 exempting the transfer of small quantities of ores, source materials and special fissile materials from the rules of the chapter on supplies (Haettu 21.11.2018.)
- 7) Council Regulation (EC) No 2964/95 of 20 December 1995 introducing registration for crude oil imports and deliveries in the Community (Haettu 21.11.2018.)
- 8) Council Regulation (EC) No 338/97 of 9 December 1996 on the protection of species of wild fauna and flora by regulating trade therein (Haettu 21.11.2018.)
- 9) Council Regulation (EC) No 601/2004 of 22 March 2004 laying down certain control measures applicable to fishing activities in the area covered by the Convention on the conservation of Antarctic marine living resources and repealing Regulations (EEC) No 3943/90, (EC) No 66/98 and (EC) No 1721/1999 (Haettu 21.11.2018.)
- 10) Regulation (EC) No 1005/2009 of the European Parliament and of the Council of 16 September 2009 on substances that deplete the ozone layer (Haettu 12.11.2018.)
- 11) Regulation (EC) No 1013/2006 of the European Parliament and of the Council of 14 June 2006 on shipments of waste (Haettu 21.11.2018.)
- 12) Regulation (EC) No 1221/2009 of the European Parliament and of the Council of 25 November 2009 on the voluntary participation by organisations in a Community eco-management and audit scheme (EMAS), repealing Regulation (EC) No 761/2001 and Commission Decisions 2001/681/EC and 2006/193/EC (Haettu 21.11.2018.)
- 13) Regulation (EC) No 1946/2003 of the European Parliament and of the Council of 15 July 2003 on transboundary movements of genetically modified organisms (Haettu 21.11.2018.)
- 14) Regulation (EC) No 648/2004 of the European Parliament and of the Council of 31 March 2004 on detergents (Haettu 21.11.2018.)
- 15) Regulation (EC) No 66/2010 of the European Parliament and of the Council of 25 November 2009 on the EU Ecolabel (Haettu 21.11.2018.)

- 16) Regulation (EC) No 713/2009 of the European Parliament and of the Council of 13 July 2009 establishing an Agency for the Cooperation of Energy Regulators (Haettu 21.11.2018.)
- 17) Regulation (EC) No 714/2009 of the European Parliament and of the Council of 13 July 2009 on conditions for access to the network for cross-border exchanges in electricity and repealing Regulation (EC) No 1228/2003 (Haettu 21.11.2018.)
- 18) Regulation (EC) No 782/2003 of the European Parliament and of the Council of 14 April 2003 on the prohibition of organotin compounds on ships (Haettu 21.11.2018.)
- 19) Regulation (EU) 2017/1938 of the European Parliament and of the Council of 25 October 2017 concerning measures to safeguard the security of gas supply and repealing Regulation (EU) No 994/2010 (Haettu 21.11.2018.)
- 20) Regulation (EU) 2017/852 of the European Parliament and of the Council of 17 May 2017 on mercury, and repealing Regulation (EC) No 1102/2008 (Haettu 21.11.2018.)
- 21) Regulation (EU) No 525/2013 of the European Parliament and of the Council of 21 May 2013 on a mechanism for monitoring and reporting greenhouse gas emissions and for reporting other information at national and Union level relevant to climate change and repealing Decision No 280/2004/EC (Haettu 12.11.2018.)
- 22) Regulation (EU) No 540/2014 of the European Parliament and of the Council of 16 April 2014 on the sound level of motor vehicles and of replacement silencing systems, and amending Directive 2007/46/EC and repealing Directive 70/157/EEC (Haettu 21.11.2018.)
- 23) Regulation (EU) No 598/2014 of the European Parliament and of the Council of 16 April 2014 on the establishment of rules and procedures with regard to the introduction of noise-related operating restrictions at Union airports within a Balanced Approach and repealing Directive 2002/30/EC (Haettu 21.11.2018.)
- 24) Regulation (EU) No 691/2011 of the European Parliament and of the Council of 6 July 2011 on European environmental economic accounts (Haettu 12.11.2018.)

Yhteensä 196 909 sanetta

Liite 2: Sanaluokkajakaumat korpuksissa

	FI-korpus	UK-korpus	EU-korpus
substantiivi	67 584	63 739	66 478
adjektiivi	12 633	10 525	14 160
verbi	30 647	27 832	29 416
sisällöllinen adverbi	4316	3730	3619
kieliopillinen adverbi	707	638	968
pronomini	3 000	4 392	3 362
partikkeli	1 123	462	811
artikkeli/tarkenne	26 171	28 099	25 297
prepositio	35 010	39 733	34 871
konjunktio	8 697	8 636	8 571
numeraali	9 059	8 542	9 290
määrittämätön	8	40	66
saneita yhteensä	198 955	196 368	196 909

Liite 3: Verrokkikorpuksista ajetus sanalistat

UK-korpuksen suuren frekvenssin sanat (≥ 1000 esiintymää) esiintymät muissa korpuksissa

sijoitus UK-korpuksesta	sana	esiintymät korpuksissa		
		UK-korpus	FI-korpus	EU-korpus
1	the	14 089	19 006	17 211
2	of	10 829	9 064	9 603
3	to	7 282	5 759	6 292
4	a	6 405	3 857	3 347
5	in	5 635	5 395	5 270
6	or	5 052	3 360	3 210
7	and	3 394	5 273	5 248
8	by	2 826	1 561	1 642
9	for	2 623	2 996	2 686
10	any	2 584	381	806
11	section	2 522	2 582	35
12	be	2 265	2 451	2 079
13	under	2 118	632	369
14	is	2 029	1 868	870
15	as	1 942	1 255	1 239
16	b	1 853	51	653
17	that	1 852	1 143	1 467
18	this	1 745	1 202	1 098
19	subsection	1 639	388	0
20	act	1 587	1 940	74
21	may	1 583	1 132	710
22	which	1 507	543	800
23	an	1 348	1 031	859
24	authority	1 177	667	502
25	with	1 131	1 040	1 796
26	on	1 100	2 449	1 415
27	shall	1 050	1 778	3 872
28	state	1 042	261	572
29	secretary	1 034	3	0

EU-korpuksen suuren frekvenssin sanat (≥ 1000 esiintymää) esiintymät muissa korpuksissa

sijoitus EU-korpuksessa	sana	esiintymät korpuksissa		
		EU-korpus	FI-korpus	UK-korpus
1	the	17 211	19 006	14 089
2	of	9 603	9 064	10 829
3	to	6 292	5 759	7 282
4	in	5 270	5 395	5 635
5	and	5 248	5 273	3 394
6	shall	3 872	1 778	1 050
7	a	3 347	3 857	6 405
8	or	3 210	3 360	5 052
9	article	2 784	55	71
10	for	2 686	2 996	2 623
11	be	2 079	2 451	2 265
12	with	1 796	1 040	1 131
13	by	1 642	1 561	2 826
14	that	1 467	1 143	1 852
15	on	1 415	2 449	1 100
16	member	1 243	83	46
17	as	1 239	1 255	1 942
18	this	1 098	1 202	1 745

Liite 4: Verrokkikorpuksista ajetut n-grammilistat

UK-korpuksen suuren frekvenssin n-grammien (≥ 200 esiintymää) esiintymät muissa korpuksissa

sijoitus UK-korpuksessa	n-grammi	esiintymät korpuksissa		
		UK-korpus	FI-korpus	EU-korpus
1	secretary of state	1 013	0	0
2	the secretary of	1 012	0	0
3	the secretary of state	997	0	0
4	in relation to	803	6	51
5	the purposes of	434	20	63
6	by virtue of	430	43	7
7	for the purposes	423	20	65
8	for the purposes of	419	19	63
9	under this section	347	3	0
10	in respect of	251	12	56
11	in accordance with	236	266	830
12	in this section	222	31	0

EU-korpuksen suuren frekvenssin n-grammien (≥ 200 esiintymää) esiintymät muissa korpuksissa

sijoitus EU-korpuksessa	n-grammi	esiintymät korpuksissa		
		EU-korpus	FI-korpus	UK-korpus
1	in accordance with	830	266	236
2	referred to in	730	1 112	37
3	pursuant to article	394	2	3
4	to in article	330	21	0
5	referred to in article	330	20	0
6	of this regulation	326	0	0
7	in accordance with the	302	113	69
8	accordance with the	302	113	69
9	the commission shall	296	0	0
10	in accordance with article	282	3	1
11	accordance with article	282	3	1
12	the competent authority	243	18	0
13	member states shall	229	0	0
14	set out in	220	62	33
15	the transmission system	216	0	0
16	to in paragraph	215	73	1
17	referred to in paragraph	215	73	1
18	transmission system operator	203	0	0

Summary in English

Common Law and Eurolect in English Translations of Finnish Law. A Corpus Study

1 Introduction

Translation plays a crucial role in international legal communication, but legal translation poses many problems not present in other realms of LSP translation. Other specialized languages do not usually differ too greatly between source and target cultures in the way that the field is conceptualized (cf. medicine or statistics), but legal language is strictly dependent on the specific legal system. Law itself is conceptualized differently depending on the legal culture, and this terminological incongruity poses a very special challenge to legal translators.

English is a particularly problematic language pertaining to legal translations: on the one hand, the status of English as a global *lingua franca* makes it the obvious choice for international legal transactions, but, on the other hand, legal English is deeply rooted in the traditions of *common law* and is, as such, not suitable for describing continental legal systems. The problems that the use of English poses do not only pertain to terminology, because the way that texts are constructed and interpreted by common law jurists also differs considerably from the continental legal tradition (see Ángeles & Llopis 2017, 24–25).

My thesis aims to gain an understanding of the English used in translations of Finnish statutes and the extent of concordance between this translated legal English and the English of EU statutes and statutes laid down by the Parliament of the United Kingdom. The objective of this study is to present an overview on the similarities and differences between the translated legal English of Finnish texts and the English of common law as well as *EU legal English*. Of particular interest is the relationship between translated Finnish law and EU legal English, since not much is known about the nature of this new variety of legal English.

The data for the study comprise 26 Finnish Acts, 24 Regulations of the European Union, and 18 Acts of the Parliament of the United Kingdom. The study employs corpus methodology to examine the data: three comparable corpora were compiled of the texts, and their differences and similarities were examined with the corpus software tools *Wordsmith* and *Sketch Engine*.

The corpora are compared globally on the macro-level (sentence length, parts of speech, type-token ratio, lexical density) as well as locally on the micro-level (wordlists, n-grams, communication of legal logic). The results of the corpus analysis are examined in light of the theory of translation universals, particularly the translation universal hypotheses of *explicitation*, *normalization*, and *simplification*. Although the nature and existence of translation universals is debated, the universal hypotheses work as a system of measuring the textual fit of translations: the hypotheses indicate whether possible differences between the corpora may arise from within the translation process or if they are more likely caused by other factors, such as institutional drafting conventions or differences between the legal systems.

The following chapter is a review of literature relevant in terms of the theoretical framework of the study. I will shed light on the central differences between common law and continental law as well as the legal English of common law and EU legal English. An overview is also presented on two problematic features of legal translation: terminological incongruity between legal systems and communication of legal reasoning. The chapter closes with an overview on translation universals. The third chapter discusses corpus methodology and the data selected for the study, and the fourth chapter reports the central results of the study. In the fifth, concluding chapter, I will discuss the results in relation to prior research.

2 Theoretical background

2.1 Legal systems

Judicial systems differ in the way that law is conceptualized, but legal systems can be divided into two globally significant legal families: common law and continental law (also known as civil law or Romano-Germanic law) (Mattila 2017, 166). Common law originally formed in England after the Norman Conquest, and law is largely formed through precedents in the common law tradition. Judicial decisions are considered pieces of legislation as much as codified law. Consequently, lower courts are bound by the precedents made in appellate courts. (Husa 2013, 303–304.) Common law jurisdictions include, for instance, the United States (excl. Louisiana), Canada (excl. Quebec), Australia, and the United Kingdom (excl. Scotland). Continental law, on the other hand, prioritizes codified law, by which areas of law are also regulated to a greater extent than in common law, and courts cannot issue judgements that are not based on a written provision (Husa 2013, 303–304). Continental law jurisdictions include, for example, Spain, Germany, France, Italy, Sweden, and Finland.

Common law and continental legal systems thus differ significantly in the way that law is categorized and conceptualized as well as which sources of law are used. According to Mattila, these differences have a historical background: common law formed as case law in English royal courts after the Norman Conquest, and continental law traditions formed based on Roman law in European universities (Mattila 2017, 167). In a word, continental law is more theoretical whereas common law is practical law formed in the king's courts.

As for the jurisdictions discussed in this thesis, Finland, belonging to the Nordic legal family, is part of the continental legal tradition. Finnish legislation is drafted in both of the national languages, Finnish and Swedish, although Swedish versions are translations from Finnish (*Lainvalmistelun prosessiopas*). The English translations of Finnish statutes are not official, that is, they are not binding and are only meant for informative purposes. As a member state of the European Union, Regulations of the EU are binding in Finland. At the time of writing, the United Kingdom is also still part of the EU. England and Wales, however, belong to the common law tradition with mixed jurisdictions in other parts of the United Kingdom. Nevertheless, the legislation laid down by the Parliament of the United Kingdom applies in Scotland and Northern Ireland, as well. After the accession of the United Kingdom to the European Union in 1973, EU law, which was previously considered to belong to the continental system due to the continental background of the founding member states, has become much more mixed and can no longer be considered to belong solely to either legal tradition. However, English has superseded French as the dominant EU language: in 2008, 72.5 percent of the texts of the European Commission were written originally in English (Felici 2015, 126), and it has also been estimated that 95 percent of the multilingual legislation of the EU is originally drafted in English (Baaij 2013, 19). English can thus be considered a source language in EU legal texts, and even in cases where a text may have been originally drafted in another language, English is one of the co-drafting languages (see further about co-drafting e.g. Šarčević 1997, 99).

2.2 Legal English

Despite being broadly adopted as a *lingua franca* in international legal transactions, English is a problematic language in international legal situations because the language stems from the common law tradition and is conceptualized differently from continental law. However, it has been suggested that a new variety of international legal English is already developing through the use of English in the European Union. (Mattila 2017, 441; Šarčević 2018, 12). The difference to the English of common law is terminological in particular, but syntactic and

textual features of EU legal English may also differ from texts produced by common law jurists. Unfortunately, very little research exists on the features of EU legal English.

The legal English used in common law jurisdictions differs greatly from general English because of the history of common law as “common” to all courts of the king. Taking into account the role of precedents in common law, statutes are only general guidelines for judges: a statute is not considered complete until judges set precedents specifying how it ought to be interpreted (Ángeles & Llopis 2017, 25). For this reason, the English of common law is general by nature in order to apply to every possible contingency, but, on the other hand, the language has to be unequivocal and precise (Mattila 2017, 422). According to Ángeles and Llopis (2017, 24), common law jurists weigh every possible meaning of each word due to the tradition of precise wording, whereas continental jurists interpret legislation teleologically, that is, considering the original intent of the legislator as part of the interpretation process. In contrast with common law, continental legislative texts are “characterized by their open texture and [...] simple, uncomplicated sentences that state policy and principles in broad terms” (Ángeles & Llopis 2017, 24). The “obsession with textualism of Anglo-American systems” (ibid.) is a major reason why using English in describing continental legal systems is problematic.

The legal English of the European Union, although scantily researched and mainly from the perspective of terminology, may eventually evolve into a legal lingua franca. The English used in EU institutions deviates from the common law tradition in that EU legal English has formed through translation, mainly from French (Robertson 2012, 1237; Biel 2015, 150). EU texts are also drafted to be translated, which likely has an effect on how texts are constructed. Additionally, a large part of English EU texts are written by non-native speakers of English (Felici 2015, 124), which may be helpful in bringing legal English further away from common law conventions.

2.3 Challenges of legal translation

Translating legal documents differs noticeably from other types of LSP translation, because where exact sciences and many other technical fields are described with the same concepts independent of culture, law is conceptualized differently in different judicial systems (Šarčević 1997, 66–67). This terminological incongruity causes misunderstandings for translators moving between legal systems. Legal terms and their conceptual content are interpreted differently depending on the legal system, and the optimistic goal of legal translation is to produce a target provision that is equivalent to the source provision in legal effect (Šarčević 1997, 112).

In addition to the terminological incongruity between legal systems, another challenge in translating legislation is the communication of legal reasoning or legal logic. Translators have historically been instructed to translate legal texts word for word, but Kinnunen says that jurists are still suspicious of translators and their level of expertise (Kinnunen 2006, 33). According to Šarčević (1997, 166), “lawyers are concerned mainly about the logical forms underlying syntactic structure, not about grammar as an end in itself”, and Kinnunen lists knowledge of legal systems and the juridical way of thinking as part of the legal translator’s competences (Kinnunen 2006, 134). Thus, the legal translator has to be able to apply legal reasoning to produce target provisions with logical structures and legal reasoning equivalent to the source provision. This includes the “if–then” structure present in most legal rules, which contain the descriptive *fact-situation* (propositional content) and the prescriptive *statement of law* (normative content) (Šarčević 1997, 137). Conjunctions and other phrases indicating an explicit “if–then” structure include, among others, *apart from*, *except*, *if*, *in so far as*, *in the absence of*, *in the case*, *in this event*, *notwithstanding*, *provided that*, *save for*, *unless*, *when*, *whenever*, and *where* (Biel 2014a, 144). The legal translator has to be able to identify and apply structures denoting legal reasoning to be able to produce a target text that is interpreted in the same way as the source text.

Pertaining to non-authoritative translations, the responsibility for interpretation may also be the reader’s. According to Kinnunen, judges who read translations of foreign statutes are expected to be familiar with the basic concepts of the foreign legal system, and translations of statutes should be considered a foreign-language text belonging to the source culture, not the target culture (Kinnunen 2006, 21). However, Kinnunen’s remark does not apply to authoritative translations, such as Swedish translations of Finnish statutes. On the other hand, Krogsgaard Vesterager claims that word-for-word translation is still preferred for authoritative translations, whereas non-authoritative, informative translations may be freer (Krogsgaard Vesterager 2017, 107). Šarčević does not see a difference between authoritative and non-authoritative translations; the translator should in any case be able to identify the normative content of a provision and formulate a target provision with the same legal effect (Šarčević 1997, 137). The role of the translator is thus not clearly defined and seems to be dependent on the function of the target text.

2.4 Translation universals

Translation universals are suggested general features of translation appearing independent of source and target languages. Chesterman (2010, 176–177) has divided potential translation universals into two groups:

- 1) S-universals, which concern differences between source and target texts; suggested S-universals include lengthening of TTs, interference, standardization, explicitation and decreased repetition.
- 2) T-universals, which concern the textual fit of TTs among texts originally written in the TL; suggested T-universals include simplification, normalization, untypical TL structures and underrepresentation of TL unique features.

The nature and existence of translation universals is debated, and research has shown evidence both for and against universals (for an overview, see Mauranen 2004, 65). Biel remarks that one should be critical of translation universals because they continue to have the status of only hypotheses, research has been contradictory or inconclusive, and the majority of research has been conducted on translations of literary texts, “which is hardly representative of and generalizable to all translators’ behaviour, especially legal translators” (Biel 2014b, 183). According to Chesterman (2004, 11), universality as such is not as much of importance as the knowledge of patterns gained with the help of the hypotheses, and Mauranen remarks that translation universals should not be considered absolute laws, but rather general or probable tendencies (Mauranen 2006, 94). As such, translation universal hypotheses could be used in comparable corpus studies as a tool for estimating whether differences found between translations and texts written originally in the target language arise from the translation process or if they are caused by external factors.

3 Methodology and data

The textual features of English translations of Finnish legislation are expected to be closer to Regulations of the European Union than Acts laid down by the Parliament of the United Kingdom. Although legislative texts of the European Union are originally written mainly in English, EU texts are often written by non-native speakers of English. Additionally, legislative texts of the EU are written to be translated, due to which it is likely that they resemble translations textually.

Corpus methodology is applied in analyzing the data, and the results are examined in light of the translation universals of explicitation, normalization and simplification in order to assess whether differences between the corpora arise from the translation process or if they are caused by other factors. The hypothesis is tested with two specific research questions:

- 1) To what extent do the textual features of translated national legislation differ from target-language supranational legislation and target-language national legislation:
 - a. at the macro-level (sentence length, parts of speech, lexical density, type-token ratio)?
 - b. at the micro-level (wordlists, n-grams, communication of legal logic)?
- 2) Do the translation universal hypotheses of *explicitation*, *normalization* and *simplification* apply to English translations of Finnish legislation?

Comparable corpora enable gaining empirical, quantitative information about textual data. Although Biel (2014a, 89) claims that generalizations and conclusions drawn based on corpora can only apply to the sample under observation and not language as a whole, corpora facilitate finding textual patterns and phenomena from large samples of text, which would manually be arduous, if not impossible. However, corpus research on translations does not replace traditional qualitative methods; as Bhatia (2004, 224) summarizes: “qualitative analysis begins where corpus linguistics ends”. As the texts examined in this study are not analyzed using qualitative methods, I will review the information obtained on the basis of the corpora about the textual fit of the English used in translations of Finnish legislation in light of translation universal hypotheses, the literature on comparative legal linguistics, and the specificities of legal translation.

The data comprise three corpora containing statutes from the fields of energy and environment law (number of words and individual texts in table 1):

- 1) Unofficial English translations of Finnish Acts (FI-corpus)
- 2) Acts of the Parliament of the United Kingdom (UK-corpus)
- 3) English versions of Regulations of the European Union (EU-corpus).

Table 1. Number of words and texts in each corpus.

	FI-corpus	UK-corpus	EU-corpus
# of words	198,955	196,368	196,909
# of texts	26	18	24

Because the structure of texts differs significantly between the legal systems, only the normative content of each statute was included in the data. In other words, appendices, preambles and other parts of the texts that do not include legal provisions were excluded from the corpora because they might skew the results of the analysis. The texts in FI-corpus are a total population sample of all English translations of Acts issued since 1994 found in the Finlex.fi legislation portal with the search terms *energy, power, environment*, forest*, waste, electric*, ja climate*; the sample was supplemented with translations classified under keywords *waste, environment, environmental impact, forestry, ja conservation of nature*. Texts for the other two corpora were picked using random sampling until the number of words of each corpora was roughly equal. Counting by the number of words, UK-corpus is only 0.12 percent larger than FI-corpus and EU-corpus 0.11 percent smaller. Due to the negligibly small difference, wordlist frequencies, for instance, have not been indicated as *per million words*. The way that P-values were calculated for all reported differences also accounts for differences in corpus size. However, UK-corpus has a smaller number of individual texts, which may affect calculations of type-token ratio, although standardized type-token ratio should not be affected.

4 Analysis

4.1 Textual fit at the macro-level

At the global level, the textual fit of the corpus consisting of translations was examined with analyses of *sentence length, parts of speech, lexical density, and type-token ratio*. The translation universal hypotheses of *explicitation, normalization and simplification* were used to examine whether the differences might arise from the translation process.

Figure 1 shows average sentence length in each corpus as well as the standard deviation of sentence length. As can be seen in the figure, long sentences as well as a high degree of deviation in length are common in each corpus. However, whereas the difference between FI-corpus and EU-corpus is rather minimal with only 0.24 words in average sentence length and 3.2 words in standard deviation ($P = 0.79$), the difference of FI-corpus to UK-corpus is 16.36 words in average sentence length and 10.67 words in standard deviation, which is also

statistically very significant ($P < 0.001$). In UK-corpus, sentences are considerably longer, more variation is found, and sentences of over a hundred words are very common, whereas FI-corpus and EU-corpus show less variation. No evidence was found supporting translation universal hypotheses; the normalization hypothesis applies somewhat to sentence length, but the difference between FI-corpus and EU-corpus is not statistically significant.

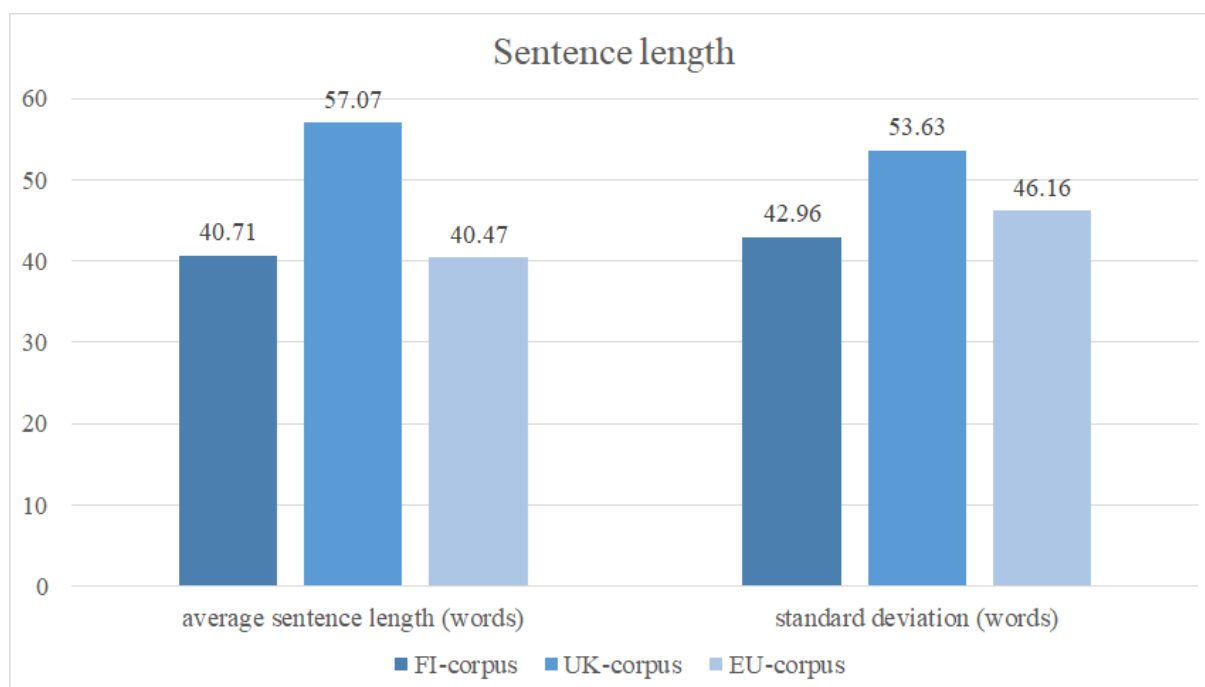


Figure 1. Average sentence lengths in the corpora

Tokens in the corpora were categorised into 11 categories according to the part-of-speech (POS) annotation given by Treetagger. A diagram of frequencies is displayed in Figure 2. The use of conjunctions does not differ significantly between the corpora, but UK-corpus has a considerably greater number of pronouns and prepositions, which does not indicate support for the explicitation hypothesis. Instead, the more frequent use of pronouns in UK-corpus and EU-corpus compared to FI-corpus as well as the more frequent use of nouns in FI-corpus may be caused by repetition, which would be an indication of simplification. However, the differences in the frequency of the use of pronouns, prepositions and nouns between FI-corpus and EU-corpus are not statistically significant, which shows that, in relation to the translation universal hypotheses, EU-corpus resembles translated language.

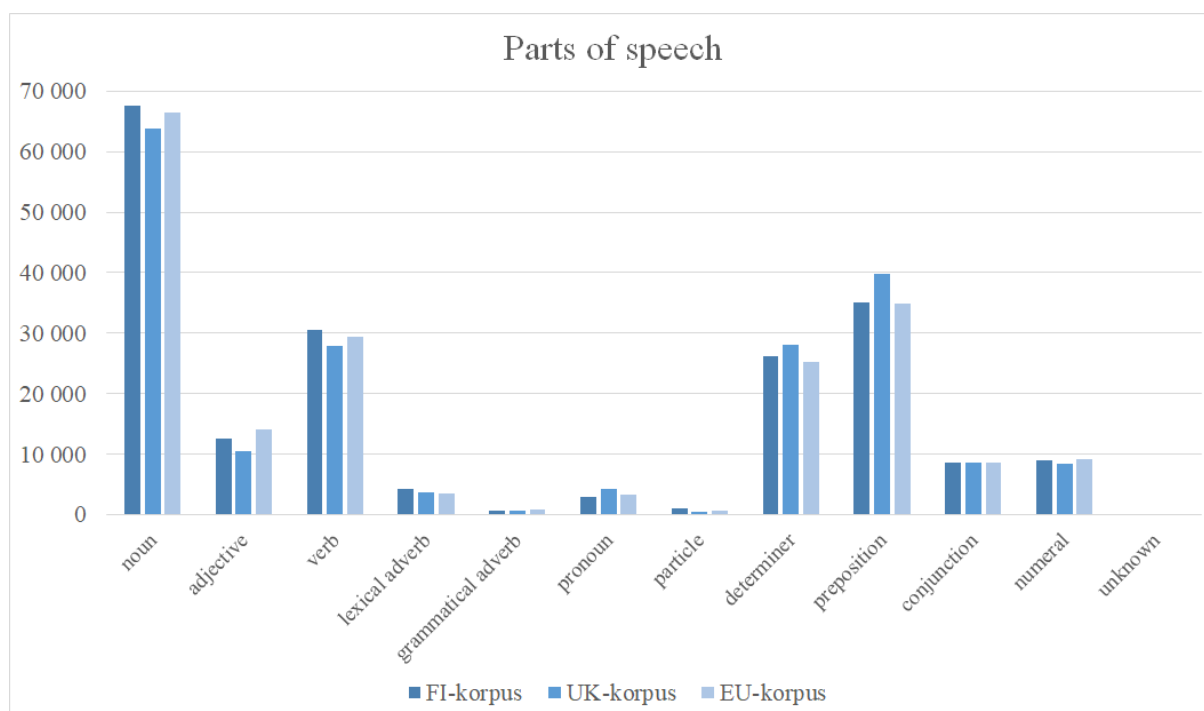


Figure 2. Frequencies of parts of speech in the corpora

Calculations of lexical density confirm this observation. Lexical density was calculated by dividing the number of lexical tokens, that is, nouns, adjectives, verbs and lexical adverbs, by the number of all tokens in the corpus. Calculated this way, FI-corpus showed a lexical density of 57.89%, UK-corpus 53.89% and EU-corpus 57.73%. The difference between FI-corpus and UK-corpus is statistically very significant ($P < 0.001$), whereas the difference between FI-corpus and EU-corpus is not ($P = 0.30$). According to the simplification hypothesis, the lexical density of translations should be lower than that of non-translation, but the results contradict this hypothesis.

Type-token ratio (TTR) was calculated by dividing the number of types with the number of all tokens in a corpus. The type-token ratio calculated from large bodies of text is inevitably small and it is also to some extent dependent on the length of individual texts (see e.g. Biel 2014a, 139), which is why *standardized type-token ratio* (STTR) calculated in intervals of 1000 words is a more apt measure of type-token ratio. TTR values of the corpora are presented in Figure 3 below.

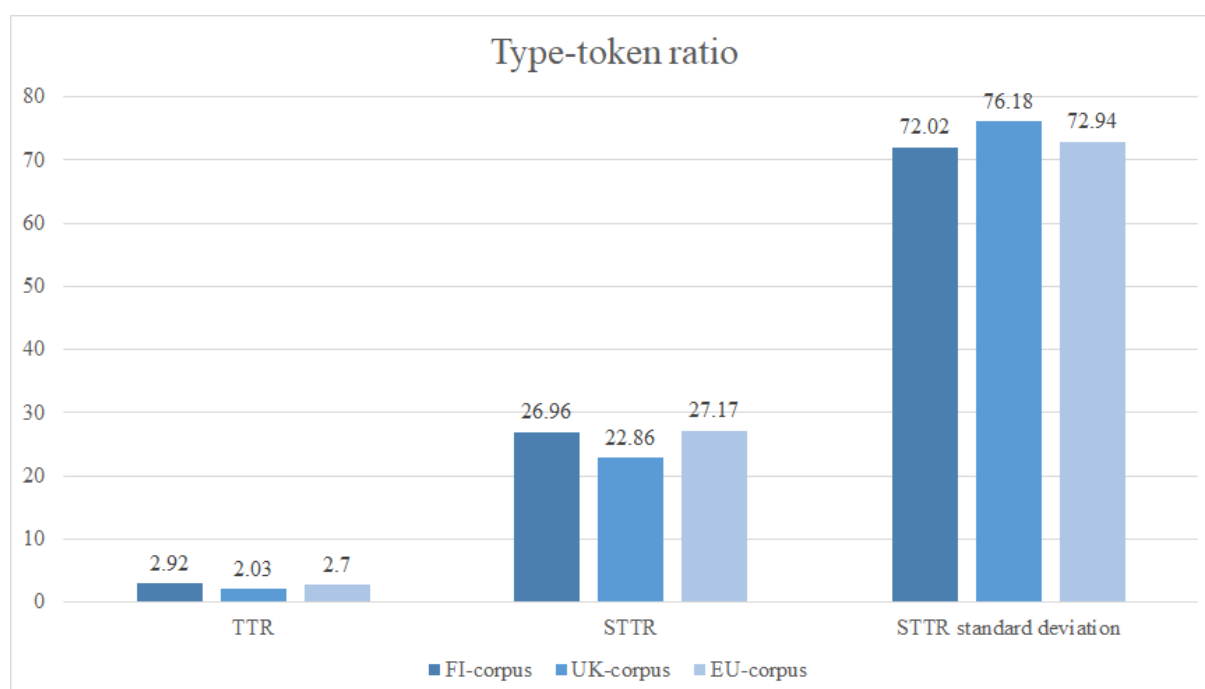


Figure 3. Type-token ratios in the corpora

As said before, UK-corpus' lower TTR value can be caused by the smaller number of individual texts, due to which the texts are also longer and may thus contain more repetition. However, the standardized type-token ratio should not be affected by this, and the STTR values show similar differences, although the differences are not statistically significant. The P-values calculated for the differences between the corpora are rather high: 0.58 for the difference between FI-corpus and UK-corpus and 0.98 for the difference between FI-corpus and EU-corpus. Although both differences are statistically insignificant, the P-values show, again, a remarkable level of similarity between FI-corpus and EU-corpus. The difference in standard deviation shows some support for the normalization hypothesis, but the level of possible normalization between the corpora is rather minimal and can be considered to be random variation due to the high P-value.

4.2 Textual fit at the micro-level

At the local level, the textual fit of FI-corpus was examined with analyses of wordlists, n-grams, and communication of legal logic. In comparison with the macro-level, much less concordance was found between FI-corpus and EU-corpus, as some of the local textual features of FI-corpus were more in line with UK-corpus.

Although FI-corpus and EU-corpus show a slight correlation in the wordlist analysis, a clear preference for the textual features of either TL corpus was not found, because P-values

calculated for differences in frequency were, for the most part, lower than 0.001. All corpora showed a strong preference for the *of* genitive instead of the possessive *s*, which is characteristic of legal language. The determiner *which* appears 1,507 times in UK-corpus, whereas only 543 and 800 times in FI-corpus and EU-corpus, respectively. It is worth noting that the frequencies of the connectives *and* and *or* are very similar in FI-corpus and EU-corpus: the P-value calculated for the difference in the use of *and* is 0.77 and 0.15 for *or*. *By* and *in* were other words that correlated between FI-corpus and EU-corpus outside of statistical significance.

Table 2. High-frequency words in FI-corpus (words appearing at least 1000 times)

wordlist position in FI-corpus	word	frequencies		
		FI-corpus	UK-corpus	EU-corpus
1	the	19 006	14 089	17 211
2	of	9 064	10 829	9 603
3	to	5 759	7 282	6 292
4	in	5 395	5 635	5 270
5	and	5 273	3 394	5 248
6	a	3 857	6 405	3 347
7	or	3 360	5 052	3 210
8	for	2 996	2 623	2 686
9	section	2 582	2 522	35
10	be	2 451	2 265	2 079
11	on	2 449	1 100	1 415
12	act	1 940	1 587	74
13	is	1 868	2 029	870
14	shall	1 778	1 050	3 872
15	by	1 561	2 826	1 642
16	as	1 255	1 942	1 239
17	this	1 202	1 745	1 098
18	referred	1 164	73	753
19	that	1 143	1 852	1 467
20	waste	1 137	243	405
21	may	1 132	1 583	710
22	provisions	1 131	286	252
23	if	1 040	568	386
24	with	1 040	1 131	1 796
25	an	1 031	1 348	859

High-frequency words in FI-corpus are presented in table 2. It was expected that FI-corpus would show a higher frequency of *which* due to clarifying paraphrases that would explain the content of terms belonging to the Finnish legal system, but the higher frequency of *which* in UK-corpus runs contrary to the explicitation hypothesis. According to Chesterman (2010, 177), a higher share of high-frequency units in translations would indicate simplification, but in this regard UK-corpus appears simpler than FI-corpus, and FI-corpus shows simplification only in

relation to EU-corpus: words appearing at least one thousand times constitute 46.47% of all tokens in UK-corpus, 40.52% in FI-corpus and 36.31% in EU-corpus.

The analysis of n-grams showed some surprising results in that high-frequency n-grams did not appear to be legal phrasemes as much as phrases used to reference other statutes and to structure the information in the texts. Additionally, a great extent of similarity was not found in the analysis of n-grams, nor evidence for or against translation universal hypotheses. Differences in statistical significance were only found pertaining to the phrases *referred to in* and *in accordance with*. *Referred to in* is used actively in FI-corpus and EU-corpus, whereas it is barely used at all in UK-corpus. The P-value calculated for the difference between FI-corpus and EU-corpus was a statistically insignificant 0.06. *In accordance with*, on the other hand, is used much more in EU-corpus than in the other the corpora, although the phrase appears more than 200 times in all three corpora. The P-value calculated for difference between FI-corpus and UK-corpus is 0,05, which is right on the verge of statistical significance/insignificance.

Communication of legal logic was analyzed with concordance searches of phrases relevant to conditional “if-then” structures as well as performatives expressing right, duty, prohibition and permission. Phrases included for the analysis of conditional structures were *apart from*, *except*, *in so far as / insofar as*, *in the absence of*, *in (the) case(s)*, *in the event*, *notwithstanding*, *provided that*, *save for*, *unless*, *when*, *whenever*, *where*, *if*, *who*, *on (the) condition*. Pertaining to the communication of right, duty, prohibition and permission, the performatives searched for included *may*, *must*, *shall*, *permit*, *forbid*, *shall not*, *must not*, and *may not*.

The analysis of conditional structures showed some support for the simplification hypothesis. The conjunction *if* is clearly overrepresented in FI-corpus, while more variation is evident in the TL corpora. *If* appears 1,040 times in FI-corpus, 568 in UK-corpus and 386 times in EU-corpus. It is worth noting that *where* is overrepresented in EU-corpus, and it appears with considerably higher frequency than *if*: *where* appears 186 times in FI-corpus, 300 times in UK-corpus and 684 times in EU-corpus. UK-corpus, on the other hand, shows considerably more variation, and the observed conditional structures in general are used less in UK-corpus, which may be caused by the tendency of common law statutes to lay down only few significant propositions of law (see Ángeles & Llopis 2017, 24).

The above-mentioned feature of common law statutes may have also had an effect on the use of performatives, as *shall* is used considerably less frequently in UK-corpus than in the other

two corpora, whereas it is overrepresented in EU-corpus: *shall* appears 1,680 times in FI-corpus, 929 times in UK-corpus and 3,673 times in EU-corpus. In relation to *shall*, FI-corpus does not appear to be closer to either TL corpora, but *may not* and *must* show similar frequencies in FI-corpus and UK-corpus: *may not* appears 88 times in FI-corpus, 82 times in UK-corpus and 13 times in EU-corpus; *must* appears 544 times in FI-corpus, 481 times in UK-corpus and 35 times in EU-corpus. The P-value calculated for the difference between FI-corpus and UK-corpus in the use of *may not* is 0.89 and 0.57 in the use of *must*, whereas the difference between FI-corpus and EU-corpus is statistically very significant ($P < 0.001$) for both *may not* and *must*.

5 Conclusions

The textual fit of English translations of Finnish statutes was found to be closer to EU legislation than legislation of the United Kingdom. The results of the corpus analysis were consistent at the macro-level, but more variation as well as some contradictory observations were evident at the micro-level of analysis. Thus, it is evident that translated Finnish law is closer to EU legal English in terms of sentence structuring. The macro-level analysis also shows a clear structural difference between EU legal English and the legal English of common law. A more terminologically-oriented approach would be of use for analyzing differences at the micro-level because of the somewhat inconclusive results of the micro-level analysis in this study.

The translation universal hypotheses of explicitation, normalization and simplification were useful in assessing textual fit, but conclusive support for the hypotheses was not found, although some evidence of normalization and simplification was reported. Explicitation could have been better analyzed with an examination of target texts in relation to source texts. The inconclusive results in terms of translation universal hypotheses indicate that legal translation differs considerably from other types of translation because the translator has to take into consideration factors specific to the legal systems between which he/she operates as well as institutional drafting conventions and guidelines.

Exactly due to the significance of institutional differences in the drafting of legal texts, this study was far from exhaustive as only one common law system was under observation. Corpora consisting of legislation from, for instance, Ireland, Scotland, Malta and Australia could be included to allow for a better comparative view of the features of legal English in different legal families (Ireland and Scotland are considered mixed jurisdictions, Malta a continental legal system, and Australia a common law system).

Nevertheless, the resemblance between the structural features of EU legal English and translated Finnish law provided valuable insight not only on English translations of Finnish law but also on the features of EU legal English itself. Considering that EU legislative texts are often drafted by non-native English speakers and to be translated into other EU languages, it is not surprising that EU legal English resembles the English of legal translations.